Emerald Books

Research in Labor Economics
Volume 30

Jobs, Training and Worker Well-Being

Solomon W. Polachek
Konstantinos Tatsiramos
Editors



JOBS, TRAINING AND WORKER WELL-BEING

RESEARCH IN LABOR ECONOMICS

Series Editor: Solomon W. Polachek

IZA Co-Editor: Konstantinos Tatsiramos

Volume 23: Accounting for Worker Well-Being
Edited by Solomon W. Polachek

Volume 24: The Economics of Immigration and Social Diversity *Edited by Solomon W. Polachek, Carmel Chiswick and Hillel Rapoport*

Volume 25: Micro-Simulation in Action
Edited by Olivier Bargain

Volume 26: Aspects of Worker Well-Being
Edited by Solomon W. Polachek and Olivier Bargain

Volume 27: Immigration: Trends, Consequences and Prospects for The United States
Edited by Barry R. Chiswick

Volume 28: Work, Earnings and Other Aspects of the Employement Relation *Edited by Solomon W. Polachek and Konstantinos Tatsiramos*

Volume 29: Ethnicity and Labor Market Outcomes
Edited by Amelie F. Constant, Konstantinos Tatsiramos and Klaus F. Zimmermann

RESEARCH IN LABOR ECONOMICS VOLUME 30

JOBS, TRAINING AND WORKER WELL-BEING

EDITED BY

SOLOMON W. POLACHEK
Binghamton University, New York

KONSTANTINOS TATSIRAMOS
IZA, Germany



United Kingdom – North America – Japan
India – Malaysia – China

Emerald Group Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2010

Copyright © 2010 Emerald Group Publishing Limited

Reprints and permission service
Contact: booksandseries@emeraldinsight.com

No part of this book may be reproduced, stored in a retrieval system, transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without either the prior written permission of the publisher or a licence permitting restricted copying issued in the UK by The Copyright Licensing Agency and in the USA by The Copyright Clearance Center. No responsibility is accepted for the accuracy of information contained in the text, illustrations or advertisements. The opinions expressed in these chapters are not necessarily those of the Editor or the publisher.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-84950-766-0
ISSN: 0147-9121 (Series)



Awarded in recognition of Emerald's production department's adherence to quality systems and processes when preparing scholarly journals for print



INVESTOR IN PEOPLE

CONTENTS

LIST OF CONTRIBUTORS *vii*

PREFACE *xi*

ON THE LINK BETWEEN INVESTMENT IN ON-THE-JOB TRAINING AND EARNINGS DISPERSION: THE CASE OF FRANCE
Audrey Dumas, Said Hanchane and Jacques Silber *1*

EMPLOYEE TRAINING AND WAGE DISPERSION: WHITE- AND BLUE-COLLAR WORKERS IN BRITAIN
Filipe Almeida-Santos, Yekaterina Chzhen and Karen Mumford *35*

INCOME INEQUALITY, INCOME MOBILITY, AND SOCIAL WELFARE FOR URBAN AND RURAL HOUSEHOLDS OF CHINA AND THE UNITED STATES
Niny Khor and John Pencavel *61*

WHY ARE JOBS DESIGNED THE WAY THEY ARE?
Michael Gibbs, Alec Levenson and Cindy Zoghi *107*

IS SENIORITY-BASED PAY USED AS A MOTIVATIONAL DEVICE? EVIDENCE FROM PLANT-LEVEL DATA
Alberto Bayo-Moriones, Jose E. Galdon-Sanchez and Maia Güell *155*

THE PROMOTION DYNAMICS OF AMERICAN EXECUTIVES
Christian Belzil and Michael Bognanno *189*

SELF-SELECTION MODELS FOR PUBLIC AND PRIVATE SECTOR JOB SATISFACTION
Simon Luechinger, Alois Stutzer and Rainer Winkelmann 233

THE SURVIVAL AND GROWTH OF ESTABLISHMENTS: DOES GENDER SEGREGATION MATTER?
Helena Persson and Gabriella Sjögren Lindquist 253

FUTILE AND EFFECTIVE WAYS TO COMBAT WAGE DISCRIMINATION
Yuval Shilony and Yossef Tobol 283

PATTERNS OF NOMINAL AND REAL WAGE RIGIDITY
Louis N. Christofides and Paris Nearchou 301

LIST OF CONTRIBUTORS

Filipe Almeida-Santos	Martifer Solar Group, Portugal
Alberto Bayo-Moriones	Universidad Pública de Navarra, Pamplona, Spain
Christian Belzil	Ecole Polytechnique, Palaiseau, France; Ecole Nationale de la Statistique et de l'Administration Economique (ENSAE), Paris, France; IZA, Bonn, Germany
Michael Bognanno	Temple University, Philadelphia, PA, USA; IZA, Bonn, Germany
Louis N. Christofides	University of Cyprus, Nicosia, Cyprus; University of Guelph, Guelph, Canada
Yekaterina Chzhen	Department of Economics and Related Studies, University of York, York, UK; Department of Social Policy, University of York, York, UK
Audrey Dumas	Laboratoire d'Economie et de Sociologie du Travail (LEST), Aix-en-Provence, France
Jose E. Galdon-Sanchez	Universidad Pública de Navarra, Pamplona, Spain; IZA, Bonn, Germany
Michael Gibbs	University of Chicago Booth School of Business, Chicago, IL, USA; IZA, Bonn, Germany

Maia Güell	University of Edinburgh, Edinburgh, UK; Universitat Pompeu Fabra, Barcelona, Spain; Center for Economic Policy Research (CEPR), London, UK; Centre for Economic Performance, London School of Economics (LSE), London, UK; IZA, Bonn, Germany
Said Hanchane	Instance Nationale de l'Evaluation, Conseil Supérieur de l'Enseignement, Rabat, Royaume du Maroc
Niny Khor	Asian Development Bank, Manila, Philippines
Alec Levenson	University of Southern California, Los Angeles, CA, USA
Gabriella Sjögren Lindquist	Swedish Institute for Social Research, Stockholm University, Stockholm, Sweden
Simon Luechinger	STICERD at London School of Economics (LSE), London, UK; ETH Zurich, Zurich, Switzerland
Karen Mumford	Department of Economic and Related Studies, University of York, York, UK; IZA, Bonn, Germany
Paris Nearchou	University of Cyprus, Nicosia, Cyprus
John Pencavel	Department of Economics, Stanford University, Stanford, California, USA
Helena Persson	Swedish Confederation of Professional Associations, Stockholm, Sweden
Yuval Shilony	Department of Economics, Bar-Ilan University, Ramat-Gan, Israel
Jacques Silber	Department of Economics, Bar-Ilan University, Ramat-Gan, Israel

Alois Stutzer	University of Basel, Basel, Switzerland; IZA, Bonn, Germany
Yossef Tobol	Iner-Disciplinary Department of Social Sciences, Bar-Ilan University, Ramat-Gan, Israel
Rainer Winkelmann	Socioeconomic Institute, University of Zurich, Zurich, Switzerland; IZA, Bonn, Germany
Cindy Zoghi	Bureau of Labor Statistics, Washington, DC, USA

PREFACE

Early models of the functional distribution of income assume constant labor productivity among all individuals. Not until human capital theory developed did scholars take into account how productivity varied across workers. According to early human capital models, this variation came about because each individual invested differently in education and training. Those acquiring greater amounts of schooling and on-the-job training earned more. However, these models neglected why one person would get training while another would not. One explanation is individual heterogeneity. Some individuals are smarter, some seek risk, some have time preferences for the future over the present, some simply are lucky by being in the right place at the right time, and some are motivated by the pay incentives of the jobs they are in. This volume contains 10 chapters, each dealing with an aspect of earnings. Of these, the first three deal directly with earnings distribution, the next four with job design and remuneration, the next two with discrimination, and the final chapter with wage rigidities in the labor market.

In a sense, analyzing earnings distribution enables one to understand human welfare, arguably the core reason for studying economics. In the first chapter, Audrey Dumas, Said Hanchane, and Jacques Silber examine an important aspect of earnings distribution. Of course, earnings vary between trained and untrained workers, but within these groups employee heterogeneity plays a role. Dumas, Hanchane, and Silber extend an approach originally introduced by Gary Fields in *Research in Labor Economics* (RLE) Volume 22 (2003) by augmenting Fields' procedure to include population subgroups. They implement the approach using recent French data. First, they find that between-group dispersion explains only 5.3% of the overall variance of earnings, implying most dispersion is within groups. From this result, they conclude that unobserved heterogeneity plays a key role in selecting those individuals that receive training. Second, they demonstrate that investment in general training affects earnings dispersion in three ways. First, training has a small direct average impact on wage inequality since its contribution to the overall variance is about 0.7%. Second, training has a much stronger effect when the training selection process is taken into account. Third, investment in training can also have an

impact on earnings dispersion via the heterogeneity of the returns to training. Based on the results, policies aimed at using vocational training to reduce wage inequality should mainly focus on better allocating training in a way that would favor women, small firms, and the less qualified workers, rather than the entire population.

Like Dumas et al., Filipe Almeida-Santos, Yekaterina Chzhen, and Karen Mumford, in the second chapter, find that returns to training vary across workers. They use 1991–2005 British Household Panel Data to explore the wage returns associated with training incidence and intensity (duration) for British employees. They find these returns differ depending on the nature of the training, the funding source for the training, the skill levels of the recipient (white or blue collar), and the age of the employee. In addition, it matters whether training was undertaken with the current or previous employer. Further, the chapter finds training to be positively associated with wage dispersion, especially for white-collar employees. As such, equal access to training programs need not reverse wage inequality, but instead exacerbate it.

Earnings dispersion across geographic regions is also important. In the third chapter, Niny Khor and John Pencavel examine income inequality, income mobility, and social welfare between rural and urban areas for the United States and China. They utilize four datasets: (1) the 1996 Chinese Household Income Project, (2) the 1991–1997 China Health and Nutrition Survey, (3) the US March 1989 and 1996 Annual Demographic Files of the Current Population Survey, and (4) the US 1994–1999 Panel Study of Income Dynamics (PSID). As a whole, China has less annual income inequality and greater income mobility than the United States. However, in contrast to the United States, they find annual income inequality in China is wider and annual income mobility is lower among rural households than among urban households. In both China and the United States, household incomes grew at a time when income inequality has widened. More importantly, using reasonable assumptions about welfare functions, the chapter finds the growth in social well-being in the United States was lower than that in China.

How jobs are designed, how compensation schemes are determined, and how workers choose their jobs are fundamental to understanding the labor market. The next four chapters deal with these issues. In the first of these chapters, Michael Gibbs, Alec Levenson, and Cindy Zoghi address a question concerning job definition. They ask: Do firms alone formulate job structures or do workers also have a substantial input? They model two kinds of job design: First are "classical" single task specialized jobs constituting division of labor. Second are "modern" multitask jobs in which workers

perform numerous aspects of the production process using at least some self-discretion. By employing a production function approach, the authors examine the trade-off between inter-task learning and gains from specialization. Their model illustrates how firm and industry characteristics explain patterns and trends in job design. Implications of the theory are tested using the 1999 BLS National Compensation Survey containing the first nationally representative sample of job characteristics. At the industry level, they find both R&D spending and computer usage to be associated with modern job design. However, particular firms tend toward extremes, choosing a modern multitask design in some establishments, and a classical single task specialized design in others. At the job level, there is a strong correlation between multitasking, discretion, skill level, and interdependence.

Given job structure, the firm still has the incentive to induce employees to maximize on-the-job effort. But as yet, there is still controversy how firms choose to motivate workers. In the next chapter, Alberto Bayo-Moriones, Jose Galdon-Sanchez, and Maia Güell test whether firms use deferred payment schemes as a motivational device. Here wages start below employee productivity, but eventually rise above it as careers progress, thus giving the employer a sanction against poor performance. What is unique about the chapter is its identification strategy. Three possibilities are considered: first, if seniority pay is used as a motivational device, then firms need not rely on other devices to monitor performance; second, if seniority pay were the result of union pressures, for example to limit management's control of the workforce, then there would be no correlation with output-based pay and monitoring; finally, if seniority pay serves as a selection device to attract applicants oriented toward long-term employment, one should observe rising average employee productivity. The authors use unique data obtained from management. They find that those firms that base their wages partly on seniority are less likely to offer explicit incentives, less likely to invest in monitoring devices, and are more likely to engage in other human resource management policies, which result in long employment relationships. They conclude that seniority-based pay is used to motivate workers.

Another question regarding job performance is how quickly workers advance through a company's hierarchy. In the next chapter, Christian Belzil and Michael Bognanno examine executive promotion. Their prime objective is to test for "fast tracks" in which workers who are promoted early on are more likely to be promoted in the future, netting out human capital, unobserved individual specific attributes, time varying firm specific variables, as well as endogenous past promotion histories. The analysis uses a 1981–1988 panel of 30,000 American executives employed in more than 300

different firms. It finds that typical easily measurable variables are relatively unimportant to predict promotion decisions once individuals get to executive levels. On the other hand, difficult–to-measure variables perhaps manifested in past promotion are important. In short, unobserved individual characteristics matter.

How jobs are designed, what they pay and how one gets a promotion do not tell the whole story. Job satisfaction is another important consideration of productivity and labor market success. In the next chapter Simon Luechinger, Alois Stutzer, and Rainer Winkelmann present an econometric methodological advance to estimate job satisfaction taking into account how workers choose their employment sector (public or private). The chapter develops a new class of ordered probit models that incorporate self-selection. The authors estimate these models using maximum likelihood techniques applied to a sample of young men from the German Socio-economic Panel. The chapter finds that workers in the public sector are better off, since they avoid the below-average job satisfaction they would have received had they chosen a private sector job.

Clearly it is in the interest of firms to motivate workers in order to maximize worker effort. But at the same time it is important to examine corporate success, such as firm survival and growth rates, as well as how government policies affect a firm's ability within the economy to behave efficiently although maintaining an equitable earnings distribution. The next two chapters examine these questions with regard to gender and race. Helena Persson and Gabriella Sjögren Lindquist do so by exploring how measures of firm performance relate to gender composition. They use a unique matched employer–employee dataset of all privately owned establishments in Sweden. To begin, they find that overall gender segregation did not change much from 1987 to 1995 and that establishment gender segregation in Sweden is comparable to that in the United States, but less than Portugal or Korea (two countries for which the authors had comparable information). With the exception of predominantly male firms, most firms become more gender integrated over time. To carry out their study, they separate new from mature establishments, and find that on average new firms are just as segregated as mature ones. However, gender-segregated firms, either male or female, have a higher risk of failing. Further, female-dominated firms have lower growth than integrated or male firms. Finally, they find that establishments that are heterogeneous with respect to gender, age, and education seem to be more successful in terms of survival and growth than more homogeneous establishments.

Their empirical results are in line with theories suggesting that heterogeneous work compositions promote higher firm payoffs. This is consistent with gains from trade coming about from comparative advantage induced by worker heterogeneity.

Workforce heterogeneity can result as a byproduct of certain antidiscrimination policies. But earnings equality is also an objective. Yet current antidiscrimination legislation can be counterproductive because of innate distortions in the way fines are collected. In the next chapter, Yuval Shilony and Yossi Tobol analyze the five major US antidiscrimination laws administered by the Equal Employment Opportunity commission: (1) The Equal Pay Act of 1963, (2) Title VII of the Civil Rights Act of 1964, (3) The Age Discrimination in Employment Act of 1967 (ADEA), (4) The Americans with Disabilities Act of 1990 (ADA), and (5) The Civil Rights Act of 1991. Each has been known to reduce rather than increase minority and female employment. Rather than finding, detecting and fining violators, as in the above laws, two alternative methods to curb discrimination are explored. One uses the tax system and the other governmental subsidies. Both result in fewer distortions than current policy.

Downward wage rigidities during the business cycle can also cause labor market distortions, especially when analyzing intertemporal changes in the wage distribution. In the final chapter, Louis Christofides and Paris Nearchou adopt a novel nonparametric approach to test for nominal and real wage rigidity. This entails examining how histograms of wage growth change from 1996–1999 based on about 11,000 collective bargaining agreements obtained for Canada. They distinguish between three regions in the wage growth distributions for which they make qualitative predictions about the nature of the distortions.

As with past volumes, we aim to focus on important issues and to maintain the highest levels of scholarship. We encourage readers who have prepared manuscripts that meet these stringent standards to submit them to RLE via the IZA website (http://www.iza.org/rle) for possible inclusion in future volumes. For insightful editorial advice, we thank Alpaslan Akay, Randall Akee, William T. Alpert, Kate Antonovics, Sowmya Wijayambal Arulampalam, Linda A. Bailey, Arnab Basu, Pieter Bevelander, Rene Boeheim, Massimiliano Bratti, Marco Caliendo, Lorenzo Cappellari, Ana Rute Cardoso, Deborah Cobb-Clark, Norma Coe, Dhaval M. Dave, Jed DeVaro, David L. Dickinson, Dimitris Georgarakos, Oliver Gürtler, Joni Hersch, David Jaeger, Martin Kahanec, Alexander Kritikos, Douglas Krupka, Astrid Kunze, Stéphanie Lluis, Corsini Lorenzo, Eduardo Melero,

Karen Mumford, Paul Oyer, Andreas Pape, Tuomas Pekkarinen, Miguel Portela, John Robst, Randolph Sloof, Murray D Smith, Arthur Van Soest, Chiara Strozzi, Nikos Theodoropoulos, Ralph Wilke, Mutlu Yuksel, Myeong-Su Yun, Anzelika Zaiceva, Zhong Zhao, and Xing Zhou.

Solomon W. Polachek
Konstantinos Tatsiramos
Editors

ON THE LINK BETWEEN INVESTMENT IN ON-THE-JOB TRAINING AND EARNINGS DISPERSION: THE CASE OF FRANCE ☆

Audrey Dumas, Said Hanchane and Jacques Silber

ABSTRACT

The aim of this chapter is to analyze the sources of earnings dispersion between trainees and nontrainees. We stress three mechanisms by which investment in general training may affect wage inequality: directly via participation to a general training program and indirectly via the selection process of trainees or the existence of heterogeneous returns on training. This chapter adopts an approach originally proposed by Fields (2003) but extends it to the breakdown of inequality by population subgroups – those

☆This chapter was started when Jacques Silber visited the Laboratoire d'Economie et de Sociologie du Travail (LEST) in Aix-en-Provence, France. A first revision was implemented when he visited the Fundación de Estudios de Economía Aplicada (FEDEA) in Madrid and the Laboratorio Riccardo Revelli at the Collegio Carlo Alberto in Moncalieri (Torino). Jacques Silber thanks these institutions for their warm hospitality.

Jobs, Training and Worker Well-Being
Research in Labor Economics, Volume 30, 1–34
Copyright © 2010 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 0147-9121/doi:10.1108/S0147-9121(2010)0000030004

who received training and those who did not. The empirical illustration is based on four French surveys, the 2006 Adult Educational Survey *and the* 2004*,* 2005*, and* 2006 Labor Force Surveys *that complement it.*

1. INTRODUCTION

Becker's (1964) theory of human capital underlines the fact that investment in human capital increases the productivity of workers and thus their wages. Like education at school, continuous vocational training is a way of increasing the human capital of individuals but there are differences between these two types of investment. What characterizes investment in training is that both employers and individuals may have an interest in investing in training. However, if a firm trains its workers, it takes the risk that its trainees may leave the firm after training and join a competitor, that is, another firm. There is thus a "poaching risk". Considering this risk, Becker (1964) stressed that a basic distinction should be made between what he called general and specific training. General training is assumed to be perfectly transferable from one firm to another, whereas specific training is supposed to increase the worker's productivity only, or at least mostly, in the training firm. For this reason Becker predicts that general training costs will be entirely borne by the individuals who are trained but they will also receive all the returns on training. On the contrary, the costs of and the returns to specific training will be shared between the employers and the individuals.

The literature on imperfect competition has, however, argued that such a distinction between general and specific training may not be relevant when markets are imperfect. The idea of these models is to assume that employees are paid less than their marginal productivity in other firms, and as a result, employers may extract a rent to finance part of the general training. The reasons for the existence of such a rent are various. First, the rent may be explained by informational asymmetries, either about the training content (Katz & Ziderman, 1990; Chang & Wang, 1996) or about the abilities of the employee (Acemoglu & Pischke, 1998). Second, minimum wages (Acemoglu & Pischke, 2003), trade unions (Booth & Chatterji, 1998), or transaction costs (Acemoglu, 1997; Acemoglu & Pischke, 1999b) can lead to a compression of the wage structure. Lastly, efficiency wages (Acemoglu & Pischke, 1999a), the guarantee of a future minimum wage (Loewenstein & Spletzer, 1998), or the heterogeneity of the firms (Stevens, 1994;

Lazear, 2003) minimize the mobility of employees, and, as a consequence, the risk of a ''poaching'' effect. This is why employers may support the costs of training, no matter how ''transferable'' the latter may be. In such a case the effects of general training may become very similar to those of specific training.

It should therefore be clear that differences in investment in training, in particular general training, may explain part of the dispersion in earnings. In fact, in order to promote economic growth and reduce wage inequality, the OECD argued in favor of investment in training, especially for those individuals who have a low level of education (OECD, 1999, 2005). For similar reasons (mainly in order to avoid under investment in training as a consequence of the ''poaching'' risk), the French government, in July 1974, made training investment compulsory for firms. French firms may thus devote a part of their wage bill to train their workers. Otherwise they have to pay a tax.

Continuous vocational training may in fact have a double impact: adapt workers to new technologies and promote social ascension for the less educated. But the reality seems to be more complex. Three reasons may explain why these previous objectives are not necessarily reached.

First, the OECD (1999) report indicates that in every country, except the Netherlands, less-educated individuals have a lower probability of being trained. This report shows that the levels of training differ significantly across (OECD) countries. Moreover, although men and women seem to have fairly equal chances of participating to job-related training, men are likely to receive greater financial support from their employers. The report also shows that training tends to fall off with age, although there are big differences between countries.[1] In the case of France, Béret and Dupray (2000) emphasized additional aspects of this unequal access to training such as the fact that training seems to be positively correlated with the professional status of the individual in the firm, the nature of his/her work contract and the size of the firm, and that it increases with seniority in the firm. Such observations may in fact lead one to conclude that the main goal of training is not to increase productivity but to ''keep the workers in the firm'' (see, Goux & Maurin, 1997). An analysis of the impact of training on the dispersion of earnings can therefore not ignore the process by which trainees are selected, since the latter may well increase wage inequality.

Second, the OECD report (1999) stressed that ''unobserved individual characteristics may determine both the probability that someone is trained and the fact that they earn higher-than-average wages after the training.'' As a consequence, part of the earnings gap between trainees and nontrainees

will be unexplained. Indeed, on the basis of the comparative study that had been conducted, this report concluded that half of the earnings gap between those who received training and those who did not is due to the fact that firms providing training pay higher salaries in any case, which is the part of earnings dispersion due to unobservables, the second half of the gap being related to factors that have a simultaneous impact on the probability of access to training and on earnings. This corresponds to that part of the dispersion of earnings that is due to the selection process of trainees.[2]

Third, the OECD (1999) study stressed that the wage premium associated with training differs between educational and gender groups, with usually higher training returns for the less-educated workers. In other words, training returns may be heterogeneous among workers and tend to reduce wage inequality. In the case of Britain, however, Almeida-Santos and Mumford (2006), who underlined that training returns may differ according to the age and occupation of trainees, found higher wages among trainees who are highly skilled employees and are over 30 years old.

There is thus room for a thorough analysis of the impact of training on the dispersion of earnings and this is precisely the goal of the present chapter.

Assume we find in a first stage that there still remains a net (of the role played by the unobserved heterogeneity) effect of on-the-job training on earnings. If we then divide our sample of workers into two groups – the first one including those who did not receive training (say, group A), the second one, those who did (group B) – we will necessarily observe that the between-group (A and B) variance of (the logarithms of) earnings is significantly different from zero. There are then two possibilities. Either the within-group variance (of the logarithms) of earnings is important, or it is not. In the latter case, this would imply that the unobserved heterogeneity that was found to have a significant impact on the probability to receive training and on the earnings themselves is in fact the "hidden" criterion for labor market segmentation. If, however. the within-group variance turns out to be important and in particular if it is much greater than the between-group variance, one would have to conclude that there is a great degree of overlapping between the two distribution of earnings, those of groups A and B. It should then be clear that the division of the sample in two groups based on a distinction between those who received and those who did not receive on-the-job training is not relevant anymore because the between-group variance turns out to be small compared to

that of the within groups. As a consequence, on-the-job training (unless the unobserved heterogeneity has also an important effect on the within-group variance) cannot be in such a case a relevant criterion of labor market segmentation.

In a second stage, we can evaluate the effects of training on the dispersion of earnings by making a distinction between these different mechanisms: the effect of training that is related to the selection process, the impact of training that is due to its heterogeneous returns, and finally its direct effect via its wage premium.

Testing such hypotheses remained a difficult task until very recently. The main goal of this chapter is to show that new developments in income inequality decomposition techniques and in the application of such techniques to regression analysis (see, Fields, 2003) allow us today to implement such tests because it has become possible to determine the exact impact of each variable not only on the overall variance of earnings but also on both the between- and within-group dispersions, the groups referring here to those who received and did not receive on-the-job training. Our study may thus shed new light on the link between training and earnings dispersion.[3]

We will proceed in three stages.

First, as has often been done in the past, in estimating an earnings function that makes a correction for the selectivity bias that is due to the selection of trainees, we will be able to check the net effect (once this selectivity bias is taken into account) of such a training on earnings.

Second, by comparing the relative importance of the between- and within-group dispersions of earnings we will find out whether there is a significant degree of overlapping between the distribution of earnings of the two groups previously mentioned (those who received and did not receive on-the-job training).

Third, by finally applying Fields' (2003) technique, we will be able to quantify the exact contribution of the observed (the explanatory) variables and of the unobserved individual characteristics to the variance of earnings. We will then be able to evaluate the contribution of the different mechanisms of training on the dispersion of earnings.

The chapter is organized as follows. In Section 2, we show how it is possible to determine the exact impact of training and other variables on the dispersion of earnings between and within groups. Section 3 describes the data sources. Our evaluation strategy is presented in Section 4. In Section 5, we present the results of our decomposition. Concluding comments are given in Section 6.

2. THE METHODOLOGY: ESTIMATING THE CONTRIBUTION OF THE EXPLANATORY VARIABLES TO THE VARIANCE OF EARNINGS

2.1. Estimating the Contribution of the Explanatory Variables to the Overall Variance

To estimate these contributions we use a recent contribution of Fields (2003) (see Appendix A).

Let us first write the earnings function as

$$y_i = \sum_{k=1}^{K+2} b_k Z_{k,i} \tag{1}$$

where y_i is the logarithm of the wage of individual i, $Z_{k,i} = X_{k,i}$, $\forall k = 1$ to K, where $X_{k,i}$ refers to the value taken by the explanatory variable k for individual i. Note that these K variables do not include the value referring to the participation (F_i) in the training program. We therefore have also $Z_{k+1,i} = F_i$ and $Z_{k+2,i} = u_i$, where u_i is the value taken by the disturbance for individual i; b_k are the average effect of the variable Z_k on the earning. Note also that we will assume below that $b_{k+1} = c$ and $b_{k+2} = 1$.

Fields (2003) has proven that standard deviation $\sigma(y_i)$ of earnings is the sum over the (K+2) variables of the product of the average effect of the variable, b_k, by the correlation between the value of the variable k and earnings, $\mathrm{Cor}(Z_{k,i}, y_i)$, and by the standard deviation of the value of the variable k, $\sigma(Z_{k,i})$:

$$\sigma(y_i) = \sum_{k=1}^{K+2} [(b_k)\mathrm{Cor}(Z_{k,i}, y_i)(\sigma(Z_{k,i}))] \tag{2}$$

The relative contribution $s_k(y_i)$ of factor k to the dispersion $\sigma(y_i)$ may therefore be expressed as

$$s_k(y_i) = \frac{[(b_k)\mathrm{Cor}(Z_{k,i}, y_i)(\sigma(Z_{k,i}))]}{\sigma(y_i)} \tag{3}$$

Expression (3) may also be written, after simplifying, as

$$s_k(y_i) = \frac{[(b_k)\mathrm{Cov}(Z_{k,i}, y_i)]}{V(y_i)} \tag{4}$$

where $V(y_i)$ denotes the variance of the logarithms of wages y_i, and $\mathrm{Cov}(Z_{k,i}, y_i)$ is the covariance of Z_k and the earnings. As a consequence, the

relative contribution of factor X_k ($k = 1$ to K) to earnings dispersion is equal to expression (4).

Similarly, the relative contribution of the participation to an on-the-job training program may be expressed as

$$s_F(y_i) = \frac{[(c)\mathrm{Cov}(F_i, y_i)]}{V(y_i)} \tag{5}$$

Finally, the relative contribution of unobserved variables (the disturbance u_i) is equal to

$$s_u(y_i) = \frac{[\mathrm{Cov}(u_i, y_i)]}{V(y_i)} \tag{6}$$

While expressions (4)–(6) give the contribution of the various explanatory factors and of the disturbance to the overall variance of the (logarithms of) wages, it is also possible to compute the contribution of these elements to the between- and within-group variance.

2.2. Contribution of the Explanatory Variables to the Between-Group Variance

When estimating the contribution of variables to the between-group variance, we stress in fact two mechanisms by which training affects the dispersion of earnings. First, we analyze what is the direct effect of training on wage inequality, net of the selectivity bias. Second, we analyze the factors that intervene in the selection process of trainees. We may thus determine the factor that increases or reduces the wage gap between trainees and nontrainees. We are also able to evaluate the impact of unobservables on the selection process and on the dispersion of earnings.

To compute the between-group variance $V_{\mathrm{BET}}(y_i)$ of the (logarithms of) earnings one has to evidently neutralize the within-group dispersion and thus assume that every worker who received on-the-job training receives the mean (logarithm of) earnings $\bar{y}_B$ of those who received such training while those who did not receive any on-the-job are assumed to receive the mean earnings $\bar{y}_A$ of those who did not receive any training (see Appendix A).

The contribution $s_{k,B}(y_i)$ of each of the (K+2) factors to the between-group variance, again using Fields' (2003) approach, will then be expressed as

$$s_{k,\mathrm{BET}}(y_i) = \frac{[(b_k)\mathrm{Cov}(\bar{Z}_k, \bar{y})]}{V_{\mathrm{BET}}(y_i)} \tag{7}$$

with $\bar{Z}_k$, the mean value of the explanatory variable Z_k in the whole population.

It is easy to show that

$$\mathrm{Cov}(\bar{Z}_k, \bar{y}) = f(1-f)(\bar{X}_{k,B} - \bar{X}_{k,A})(\bar{y}_B - \bar{y}_A) \tag{8}$$

with $\bar{X}_{k,B}$, the mean value of the explanatory variable X_k in the population of those who receive training B and $\bar{X}_{k,A}$, the mean value of the explanatory variable X_k in the population of those who do not receive training A; and that

$$V_{\mathrm{BET}} = f(1-f)(\bar{y}_B - \bar{y}_A)^2 \tag{9}$$

We may now combine expressions (7), (8), and (9) to derive that

$$s_{k,\mathrm{BET}}(y_i) = \frac{(b_k)(\bar{X}_{k,B} - \bar{X}_{k,A})}{(\bar{y}_B - \bar{y}_A)} \tag{10}$$

For the contribution of the variable F_i to the between-group dispersion, one will similarly obtain, remembering that in this case $X_{k,MB} = 1$ and $X_{k,MA} = 0$,

$$s_{F,\mathrm{BET}}(y_i) = \frac{(c)}{(\bar{y}_B - \bar{y}_A)} \tag{11}$$

Finally, the contribution of the disturbances to the between-group dispersion will be written as

$$s_{u,\mathrm{BET}}(y_i) = \frac{(\bar{u}_B - \bar{u}_A)}{(\bar{y}_B - \bar{y}_A)} \tag{12}$$

where $\bar{u}_B$ and $\bar{u}_A$ are, respectively, the mean values of the disturbances in groups B and A.

It is then easy to show that the sum of all the contributions to the between-group dispersion is equal to 1.

2.3. Contribution of the Explanatory Variables to the Within-Groups Variance

By comparing the contributions of factors to the dispersion of earnings separately for the subpopulations of trainees and nontrainees, we can see to which extend the returns to training may be heterogeneous among trainees and contribute to an increase or a reduction in the dispersion of earnings. The returns to training may be heterogeneous for two reasons: first training

may have a higher effect on the productivity of some workers, depending on their individual characteristics. Second, since there are different kinds of training programs, each program can have different return on training. Moreover, one may assume that training programs are not allocated randomly among the workers.

As is well known, the within-group variance is equal to the weighted sum of the variance within each of the two groups, A, $V_A(y_i)$ and B, $V_B(y_i)$, the weights being the population shares (f and $(1-f)$) of the two groups, so that the contribution $s_{k,\mathrm{WITH}}(y_i)$ of each of the $(K+1)$ factors[4] to the within-group variance may then be written as

$$s_{k,\mathrm{WITH}}(y_i) = (1-f)\frac{[(b_k)\mathrm{Cov}(Z_{k,i\in B}, y_{i,i\in B})]}{V_B(y_i)} + (f)\frac{[(b_k)\mathrm{Cov}(Z_{k,i\in A}, y_{i,i\in A})]}{V_A(y_i)} \tag{13}$$

We may therefore conclude, using all the previous results, that the contribution of a given factor k ($k = 1$ to K) to the total variance V_{TOT} of the logarithms of wages is the sum of three elements:

– its impact via its contribution to the within group A variance V_A; this effect will be expressed as

$$\left[\frac{(b_k)\mathrm{Cov}(Z_{k,i\in A}, y_{i\in A})}{V_A(y_i)}\right]\left[\frac{(f)V_A(y_i)}{V_{\mathrm{WITH}}(y_i)}\right]\left[\frac{V_{\mathrm{WITH}}(y_i)}{V_{\mathrm{TOT}}(y_i)}\right] = (f)\left[\frac{(b_k)\mathrm{Cov}(Z_{k,i\in A}, y_{i\in A})}{V_{\mathrm{TOT}}(y_i)}\right] \tag{14}$$

– its impact via its contribution to the within group B variance V_B; this effect will be expressed as

$$\left[\frac{(b_k)\mathrm{Cov}(Z_{k,i\in B}, y_{i\in B})}{V_B(y_i)}\right]\left[\frac{(1-f)V_B(y_i)}{V_{\mathrm{WITH}}(y_i)}\right]\left[\frac{V_{\mathrm{WITH}}(y_i)}{V_{\mathrm{TOT}}(y_i)}\right] = (1-f)\left[\frac{(b_k)\mathrm{Cov}(Z_{k,i\in B}, y_{i\in B})}{V_{\mathrm{TOT}}(y_i)}\right] \tag{15}$$

– its impact via the between-group variance V_{BET}; this effect will be expressed as

$$\left[\frac{(b_k)\mathrm{Cov}(\bar{Z}_k, \bar{y})}{V_{\mathrm{BET}}(y_i)}\right]\left[\frac{V_{\mathrm{BET}}(y_i)}{V_{\mathrm{TOT}}(y_i)}\right] = \left[\frac{(b_k)\mathrm{Cov}(\bar{Z}_k, \bar{y})}{V_{\mathrm{TOT}}(y_i)}\right] \tag{16}$$

We thus end up with a total impact of the variable k expressed as

$$\frac{[(f)(b_k)\mathrm{Cov}(Z_{k,i\in A}, y_{i\in A})] + [(1-f)(b_k)\mathrm{Cov}(Z_{k,i\in B}, y_{i\in B})] + [(b_k)\mathrm{Cov}(\bar{Z}_k, \bar{y})]}{V_{\mathrm{TOT}}(y_i)} \tag{17}$$

Similar results may be derived for the contribution of the variable on-the-job training F_i, and of the disturbances.

3. THE DATA SOURCES

3.1. The Samples

We consider four French datasets: the *2006 Adult Education Survey* and the *2004*, *2005*, and *2006 Labor Force Surveys*.

In the *2006 Labor Force Survey,* individuals were interviewed about their employment and wages situation in the first, second, third, and fourth trimester of 2006. The *2006 Adult Education Survey* is a survey that complements the *2006 Labor Force Survey*. All the individuals who were interviewed in 2006 had also to indicate whether they had participated in training programs during the past 12 months and, if so, describe the context and the type of training programs. We consider individuals who were interviewed in 2006 both for the *2006 Labor Force Survey* and the *2006 Adult Education Survey*. These surveys provided us with information on their training participation and their professional situation after training. To get information before their participation to a training program, we consider individuals that were also interviewed six quarters earlier for the *2004* or *2005 Labor Force Surveys*.

We restrict our sample to have a more homogeneous population. We only consider individuals who work in the private sector. We exclude workers from the energy sector because the latter includes too many workers who work in the same firm, and workers from the farming sector because their characteristics may be too specific.

We also delete individuals who work in firms that have less than 10 workers. As mentioned previously in France, firms are compelled to train their workers but the rate of compulsory training investment represents 0.55% of the wage bill for firms with less than 10 workers, 1.05% for firms with 10–20 workers, and 1.6% for firms with more than 20 workers.

We have thus deleted workers from the smallest firms, which may well have very different policies as far as investment in training is concerned.

We have also limited our analysis to vocational continuous training. We have thus excluded any training that is conducted partially in the firm and partially at school. We have also excluded training programs that do not have a professional purpose.

Finally, we have exclusively considered training programs that are assumed to be useful to other sectors than that of the training firm. We have thus assumed that the training programs that are the object of our analysis are ''general'' in the sense of Becker. Two reasons motivated this choice. First, Becker's model stresses the fact that general training has a higher impact than specific training on the wages of workers. As a consequence, general training may have a higher impact on wage inequality than specific training. Secondly, in order to better control the selectivity bias, it is important to have a more homogeneous definition of training, because the selection of trainees varies with the characteristics of training programs.

As a result, we ended up with a sample of 2,966 individuals, and 28.2% of the individuals in this sample participated in a training program.

3.2. Summary Statistics

As far as earnings are concerned, they refer to monthly wages measured in euros and include bonuses. The mean wage of the sample is 1,557 euros per month. Table 1 indicates that trainees earn significantly more than the average since their mean wage is 1,827 euros. This results seems to confirm

Table 1. Wages Statistics for the Sample and for the Trainees.

Wage	All ($n = 2{,}966$)			Trainees ($n = 837$)		
	Mean	Median	SD	Mean	Median	SD
Wages in 2006 (in euros)***	1,625.3	1,425	892.2	1,943.3	1,667	1,003
Wages 6 trimesters earlier (in euros)***	1,557.1	1,375	870.6	1,827	1,582	957.3
Wage growth (in %)*	8.9	3.7	31.7	10.6	5.2	31.3

Note: Significant differences between trainees and nontrainees:
*10% level of significance.
**5% level of significance.
***1% level of significance.

that participation in a training program has an impact on wages, as predicted by Becker's model. It turns out, however, that trainees have also a significantly higher wage before they receive training. The participation to a training program is therefore not random and it is generally the better-paid workers that are trained. To better identify the impact of training on wages, we can compare the wage growth of trainees and nontrainees. The statistics on wage growth confirm that the latter is higher (at a 10% significance level) among workers who participate in a training program. Training may therefore increase the dispersion of earnings because of its direct effect on wage as well as the selection process. We may also notice that the dispersion of wages is also higher than average in the subgroup of trainees. In fact, it seems that the returns to training or the selection process of trainees are heterogeneous, depending on the form of training programs.

Since we observed that trainees earn more than nontrainees even before their training, because of the selection process, we can check to which extent trainees and nontrainees have different characteristics and then see the factors that have a significant influence on the access to training. Table 2 thus shows that the proportion of women and foreigners is significantly smaller among trainees. Trainees are also significantly older, but their seniority in the firm is not significantly different from that of nontrainees. It also appears that the level of education is significantly higher among trainees and this is also true for the level of qualification of the job. There are also differences in the job position between the two groups. Also, trainees usually come from significantly larger firms and the sector in which they work is different from that of nontrainees. The firms in which trainees work have also more frequently introduced new equipments or new work organizations during the past 12 months. Finally, Table 2 shows that trainees are generally individuals who work more than 40 hours per week, even though the regular weekly duration of work is 35 hours in France. This suggests that the proportion of executives and managers is highly represented in the group of trainees.

All these observations tend to confirm that the selection process of the trainees is likely to increase the wage gap between trainees and nontrainees.

4. THE EVALUATION STRATEGY

4.1. The Econometric Method

To estimate the contribution of training participation to the dispersion of earnings, we need to correctly estimate the parameter c, and, as a

Table 2. Characteristics of the Individuals in the Sample and of the Trainees.

Variables	All	Trainees
Women**	46.3%	43.4%
Foreigners***	5.5%	2.9%
Mean number of children**	1.08%	1.16%
Mean seniority (in months)	151.9	146.9
Mean age (in year)***	41.6	39.8
Full time: more than 40 hours	25.9%	33.7%
Full time: between 35 and 39 hours	55.4%	51.6%
Permanent contract*	92.7%	94.6%
Diploma***		
None	25.8%	12.3%
BEPC (former examination at the end of the first stage of secondary education)	38.3%	33.33%
Baccalaureate	14.2%	16.5%
BTS DUT (two years after baccalaureate in technological studies)	9%	15.1%
DEUG ((two years after baccalaureate in university)	1%	1.9%
Studies in paramedical field	1.3%	2.2%
Grande école (higher education institution with competitive entrance examination)	3.4%	7.1%
Master or doctorate degree	7%	11.7%
Work schedule		
Working on Sunday**	19.8%	22.5%
Working on Saturday***	42.5%	47.1%
Working in the evening	15.3%	16.5%
Working at night	15.1%	16.3%
Flexible schedule*	29.3%	31.9%
Working at home**	10.1%	12.2%
Position***		
Production, manufacture	26.9%	21.5%
Repairing, cleaning	7.3%	7.7%
Hygiene and security	8.2%	2.9%
Transport	7.6%	5.9%
Secretarial	6.5%	7.2%
Administration	7.9%	10.8%
Trade	10.9%	12.5%
Research	6.7%	12%
Teaching	8.5%	10%
Level of qualification of the job***		
Unskilled worker	11.2%	6.7%
Skilled worker	23.4%	14.5%
Employee	31.3%	26.6%

Table 2. (*Continued*)

Variables	All	Trainees
Technician, supervisor	15%	23.2%
Engineer, executive	14.4%	24.4%
Director, manager	0.6%	1.3%
Other	3%	2.6%
Sector***		
Industry	31.2%	35.6%
Construction	6.6%	5.1%
Trade and repairs	13.8%	13.4%
Education, health, and social actions	12.9%	14.7%
Services	35.4%	31.2%
Firm size***		
10–49 workers	33.4%	30.7%
50–199 workers	25.5%	28.1%
200–499 workers	13.7%	18.3%
Larger than 500 workers	13.1%	16.9%
Unknown	14.3%	6.1%
Paris	5.5%	5.6%
Trimester of interview***		
First	20.7%	22.6%
Second	26.3%	23.7%
Third	26.2%	22.5%
Fourth	26.8%	31.3%
Mean unemployment rate	9.1%	9.1%
Changes		
New equipment***	29.7%	48.8%
New work organization***	32%	48.9%
New job: tenure less than one year and a half	9.8%	9.3%
Higher qualification job	2.6%	3%
Smaller qualification job	1.7%	1.5%
Higher working hours**	5.2%	6.5%
Smaller working hours	4.6%	5.4%
Positions different	5.1%	5.6%
Employment contract different**	6.2%	4.4%

Note: Significant differences between trainees and nontrainees:
*10% level of significance.
**5% level of significance.
***1% level of significance.

consequence, to control for the selection process that occurs as far as access to training is concerned. The issue here is that the average effect c of training on wages is likely to be biased because there are factors affecting both the access to training and wages.

A first strategy to control for this selectivity bias is to introduce in the wage equation all the observable variables that affect both the access to training and wages.[5] We can distinguish several groups of control variables, depending on the labor market theory[6] one chooses. According to human capital theory, the following individual characteristics should be considered: gender, nationality, number of children and its square, age and its square, and dummies corresponding to the educational level. According to models stressing the idea of internal market and job matching, the following variables should be included: seniority and its square, weekly duration of work (more than 40 hours, 35–39 hours, 30–34 hours, 15–29 hours, less than 15 hours), type of contract (permanent, temporary, other), qualification level of the job (unskilled worker, skilled worker, employee, technician or supervisor, engineer or executives, director or manager, unknown), position (production or manufacture, repairing or cleaning, hygiene or security, transport, secretarial, administration, trade, research, teaching), type of work schedule (Sunday, Saturday, night, evening, at home, flexible hours). According to the theory of segmented markets, the size of the firm (10–49 workers, 50–199 workers, 200–499 workers, more than 500 workers, unknown), the sector (industry, construction, trade and repairs, education, health and social actions, services) and the region in which the firm is located (Paris or not) should be taken into account. Finally, more recent theories stressing imperfect competition would recommend introducing the unemployment rate in the area (the French "departments") in which the firm is located, in order to control for the degree of competition in the labor market. All these variables are summarized under the label X. We also control for changes (W) that may have affected the situation of the worker in the 18-month period preceding the date at which the survey took place. We thus control for changes in the schedule of work, in the employment contract, the level of qualification of the job, the position of the worker, and the firm. We also control for changes in the situation of the firm in which the worker is employed. We introduce variables indicating whether a new organizational framework (in the department or in the team in which the individual works), a new equipment, or a new production technique was introduced. Finally, we also control for the trimester in which the interview took place. In fact, Table 1 indicates that this trimester is significantly different for trainees and nontrainees and may thus be a source of bias.

We estimated Model (1) with OLS

$$y_{i,2006} = cF + b_1 X_i + b_2 W_{i,t} + u_{i,t} \qquad \text{(Model 1)}$$

There may, however, be also a selection bias due to unobserved variables that were not considered in the model. These unobserved variables may reflect unobserved individual characteristics such as abilities and motivation or unobserved characteristics of the firm such as its training or wage policies. We could apply methods of instrumental variables or Heckman's two-step procedure (1979) to control for unobserved heterogeneity. But these methods require instruments that must respect two criteria: they have to predict training participation but not the workers' wages. It is extremely difficult to find such variables in the case of training, and the estimation depends strongly on the relevance of instruments. This is why we have excluded this type of strategy.

If we assume, in a first stage that the unobserved heterogeneity is time invariant, we may choose another approach, namely apply the method of first differences to estimate the average effect of training, net of selectivity bias. In such a case the time-invariant unobserved heterogeneity is deleted thanks to first differences, as are all time invariant control variables, X. The dependent variable becomes the wage growth Δy_i during the 18 months preceding the interview. We then apply OLS to Model (2) and derive

$$\Delta y_i = cF + \delta W_{i,t} + u_{i,t} \qquad \text{(Model 2)}$$

If, however, the unobserved heterogeneity is time variant, the previous method does not control all types of selectivity biases. In such a case we suggest to include all the control variables in the model in differences, even those that are time invariant, because they proxy the unobserved time-variant unobserved heterogeneity. Thus the selectivity bias may be due to differences in training investment or firm performances and they can be captured via the size of the firm and the sector. Similarly, differences in learning and in the motivation of the workers can be controlled by individual and job characteristics. We then apply OLS to Model (3). The average training effect c is then assumed to be well estimated via Model (3).

$$\Delta y_i = cF + \delta_1 X_{i,t} + \delta_2 W_{i,t} + u_{i,t} \qquad \text{(Model 3)}$$

If, however, we want to analyze the contributions of control variables to the variance of earnings, we need to derive the impact of the variables on the wage level and not on the wage growth. To get such unbiased estimates of the impact of training, we propose an alternative method, one that is a mix of the OLS and first differences methods. We consider Model (1) but

introduce dummies Q that indicate the position of the individuals in the wage distribution before training. In other words, we introduce variables indicating whether the worker is ranked in the first decile, in the range between the first decile and the first quartile, in that between the first quartile and the median, in that between the median and the third quartile, in that between the third quartile and the ninth decile, and finally in the last decile.

$$y_{i,2006} = cF + b_1 X_i + b_2 W_{i,t} + b_3 Q_i + u_{i,t} \qquad \text{(Model 4)}$$

We apply OLS to Model 4 and can then check whether the average training effect c is similar in Models 3 and 4 and hence whether the selectivity bias has been neutralized. We present the results of the estimations of these various models in the following section.

4.2. Estimating Wage Regressions

Estimates of Models (1), (2), (3), and (4) are summarized in Appendix B.

It appears that the participation to a general training program has a significant impact on wages whatever the model considered. The estimates of the average training effect in Models 3 and 4 are very close, 3.61% for Model 3 and 3.76% for Model 4. The difference may be explained by measurement errors in training participation. Indeed, some workers may not remember whether they participated in a training program during the past 12 months, and that may imply an underestimation of the impact of training that would be higher in a first difference model (Freeman, 1984). As a result, we can assume that the impact of training on earnings "net of the selectivity effect" is approximately 3.76%. This result shows that general training returns are sufficiently high to have an impact on earnings and justifies policies aiming at promoting training investment in order to modify the distribution of wages.

As far as the parameters of the other explanatory variables are concerned, there are of the same sign in Models 4 and 1, even though they are smaller in Model 4 because part of the unobserved heterogeneity is controlled for. Let us take a closer look at the results of Model 4.

As expected, women get a wage, which, ceteris paribus, is lower than that of men (4.9%). Earnings rise with the level of human capital. One may thus observe that those who have a higher education diploma (Grande école, Master or Doctorate), ceteris paribus, earn 16% more than those who have no diploma. A baccalaureate increases wages by 7.5%. As far as the socio-professional category is concerned, executives, engineers, and individuals

who are part of the managerial staff earn significantly more than the other categories. Also, note that job security seems to play a discriminating role – those having a temporary work contract earn less, ceteris paribus, than those having a permanent contract. Seniority has, as expected, a nonlinear effect but note its weak impact. We do not find a significant effect of age and even sometimes observe the contrary.[7]

It also appears that larger firms offer higher wages, but among sectors wage differences are rather small if we ignore the sectors of education, health, and social work. As far as positions are concerned, only monitoring and cleaning have smaller wages than a production position. Firms located in Paris offer higher wages (8.6%). Several variables have, however, no significant impact on wages, this being true for the type of contract, the work schedule of the job, the introduction of a new working organization or of new equipment in the firm. Professional changes between the dates of the two interviews do not affect the wages in 2006 except when the working hours increase. Finally, the trimester in which the survey took place does not have any significant effect on wages.

As a whole, the results obtained when estimating earnings functions are the ones we expected. Of particular interest is the fact that training has a net effect on earnings. This therefore allows us to decompose the variance of earnings as a function of participation in training.

5. THE RESULTS: DECOMPOSING THE VARIANCE OF EARNINGS

Table 3 gives the decomposition of the total variance of the logarithm of wages into two components, the between- and the within-group variances (the groups being those who received training and those who did not).

Table 3. Decomposition of the Total Variance of the Logarithm of Wages.

	Absolute Value	%
Total variance	0.3025	100
Between variance	0.0159	5.3
Within variance	0.2866	94.7
Within trainees variance	0.2300	21,4
Within nontrainees variance	0.3089	73,3

It appears that most of the dispersion (94.7% of the variance) takes place within groups while the between-group variance represents only 5.3% of the total variance.

We can thus see that there is a high degree of overlapping between the wage distributions of trainees and nontrainees. Let us first analyze the determinants of the earnings gap between the two subgroups and then focus on the factors of inequality within each group. We will then be able to conclude to which extent training investment affects the dispersion of earnings.

5.1. Contributions of the Various Variables to the Between-Group Variance

In a first step, to interpret the contributions of variables to the between groups variance, we sum for each variable the contributions of its modalities. This is indicated in bold letters in Table 4. It appears that working hours are the most important determinant of wage dispersion (13.6%). We should remember here that the dependent variable refers to monthly earnings (because we had no way of estimating hourly wages) so that the role played by the number of hours of work should not be surprising.

Once working hours are taken into account, the participation to a general training program explains 13.7% of the earnings gap between trainees and nontrainees. This result confirms that most of the between-group wage inequality is not directly related to returns on training. In fact, it is rather the process by which trainees are selected that leads to a between-group inequality in wages. Indeed, the between-group variance is explained by working hours (13.6%), participation to training (13.7%), and individual characteristics ($100\% - 13.7\% - 13.6\% = 72.7\%$).

Let us now check whether the process by which trainees are selected tends to increase or reduce the between-group variance in wages. The two main variables that contribute significantly to the between-group variance of the (logarithms of) earnings are the level of qualification of the job (13.1%) and the educational level of the individual (11.5%). Other relevant variables are the size of the firm (4.6%), the position of the worker (3%), the introduction of new equipment (1.7%), and the gender of the worker (0.7%). As stressed previously, these variables correspond also to the criteria on the basis of which trainees are selected and they tend to increase inequality. On the contrary, seniority (−2.2%) and the sector in which the individual works (−1.1%) are criteria of selection that tend to reduce the wage gap between trainees and nontrainees. Finally, note that the residuals do not have any

Table 4. Decomposition of the between Variance of the Logarithm of Wages.

Variable	Contribution (%)	SE	Elasticity
Training access	13.407	(3.667***)	0.2805
Trimester	0.248	(0.649)	
Children	0.217	(0.261)	
Square of number of children	−0.511	(0.381**)	
Women	0.717	(0.429**)	−6.89
Foreigner	−0.241	(0.320)	
Age and seniority	−2.232	(0.975***)	
Seniority	−1.554	(1.252*)	
Square of seniority	1.326	(0.933*)	
Education level	11.529	(2.020***)	
BEPC	−0.599	(0.366**)	−4.081
Baccalaureate	0.857	(0.478**)	8.779
Technological studies	3.087	(0.848***)	3.345
Paramedical studies	0.664	(0.343**)	24.093
Grande école	3.68	(0.968***)	5.577
Master and doctorate	3.77	(0.982***)	4.257
Working hours	13.649	(2.051***)	
Work 35–39 hours	0.772	(0.390***)	−5.443
Work 15–30 hours	3.673	(1.214***)	−8.484
Less 15 hours	8.835	(1.871***)	−11.561
Higher working hours	0.462	(0.308**)	15.569
Job Contract	0.243	(0.473)	
Job schedule	−0.024	(0.439)	
Qualification level	13.109	(1.941***)	
Qualified worker	−1.648	(0.766***)	−2.251
Employees	−0.827	(0.600*)	−4.303
Technicians, supervisor	3.233	(0.972***)	2.473
Engineer, executives	11.094	(1.878***)	2.025
Director, Manager	1.264	(0.629***)	28.463
Position	2.953	(1.235**)	
Hygiene and security	1.755	(0.842**)	−3.805
Teaching	0.466	(0.397*)	13.37
Sector	−1.058	(0.68*)	
Education, health, social action	−0.889	(0.582**)	11.296
Paris	0.051	(0.306)	
Firm size	4.550	(1.022***)	
10–49 workers	0.453	(0.327**)	−7.531
50–199 workers	−0.357	(0.276**)	7.883
200–499 workers	−0.578	(0.397*)	4.418
Firm size unknown	5.032	(1.083***)	−2.455
New equipment	1.678	(0.994*)	1.059
New working organization	−0.325	(0.871)	
New job	−0.072	(0.188)	

Table 4. (*Continued*)

Variable	Contribution (%)	SE	Elasticity
Dummies of position in wage distribution	42.403	(2.539***)	
1st deciles	25.141	(3.198***)	−3.871
1st deciles to 1st quartile	15.602	(3.189***)	−4.228
1st quartile to 2nd quartile	12.411	(3.314***)	−4.389
2nd quartile to 3rd quartile	−4.843	(2.815*)	8.538
3rd quartile to 9th deciles	−5.908	(1.379***)	3.733
Residual	9.89E−10	(1.21E−09*)	

Note: Standard errors were derived from a bootstrap procedure with 1,000 replications. Significant differences between trainees and nontrainees:
*10% level of significance.
**5% level of significance.
***1% level of significance.

significant impact on the dispersion of wages and this seems to confirm that the selectivity bias is well controlled.

All these results imply that the process by which trainees are selected worsen the initial educational and social inequalities between trainees and nontrainees. Similarly, the discrimination against women in terms of wages becomes stronger because of unequal access in training. Finally, the process by which trainees are selected seem to reinforce internal and segmented markets.

Finally, note that approximately 42% of the between-group variance of earnings is explained by unobserved heterogeneity, the latter being captured, as explained previously, by the set of dummies indicating the position of the individuals in the wage distribution, before the training took place (42.5%). One may observe the important role played by individuals belonging to the first decile (25.1%), that is, those who are the least paid. We tend to believe that this unobserved heterogeneity may actually represent the accumulation of discriminating characteristics. In other words, it may be much more difficult to have access to a training program when one is at the same time, a woman, an unskilled worker, employed by a small firm, and working part time. As our model is linear, the impact of interactions between the characteristics is not taken into consideration and may then be captured by the set of dummy indicators that were introduced.

Some additional intuitive interpretation of the results may be derived from Eq. (10) which was expressed as

$$s_{k,\mathrm{BET}}(y_i) = \frac{(b_k)(\bar{X}_{k,B} - \bar{X}_{k,A})}{(\bar{y}_B - \bar{y}_A)} \tag{10}$$

Let us now rewrite (10) as

$$\frac{(\bar{y}_B - \bar{y}_A)}{(\bar{X}_{k,B} - \bar{X}_{k,A})} = \frac{(b_k)}{s_{k,\mathrm{BET}}(y_i)} \tag{18}$$

Since the variables $\bar{y}_B$ and $\bar{y}_A$ are logarithms, their differences may be interpreted as the percentage difference in the average earnings of the two groups. Moreover, for all the dummy variables, the expression $(\bar{X}_{k,B} - \bar{X}_{k,A})$ refers in fact to the difference between the percentage of individuals in group A who have characteristic k (e.g. are "women") and the corresponding percentage in group B. Therefore in this case, the ratio $(b_k)/(s_{k,\mathrm{BET}}(y_i))$ is a kind of elasticity and it shows by how much the percentage difference between the average earnings in the two groups will increase (in absolute, not relative terms) when the gap between the percentages of individuals who have characteristic k in the two groups increases by 1% (here also in absolute, not relative terms). Let us see what this implies for the five variables mentioned previously, by looking at the data of Tables 2 and 4 and at Appendix B. Table 2 indicates for example that 43.4% of those who receive training are women while the corresponding percentage among those who do not receive training is 47.4%. The difference between these two percentages is hence equal to 4%. Remembering that the difference between the average values of the logarithms of earnings in the two groups is equal to 28.05%, we derive that the ratio $(b_k)/(s_{k,\mathrm{BET}}(y_i))$ for this variable is $-0.0494/0.7173 = -6.89$. In other words, assume that this gap between the two groups in the percentage of those being women decreases by 1%, from 4% to 3%. This then implies that the average gap in earnings between the two groups will decrease by 6.9%, from 28.1% to 21.2%.

When we analyze elasticities, we observe that the net effect of training has a small impact on the between-group variance, because the corresponding elasticity is 0.281. This number implies that if the proportion of trainees increases by 1% (from 28.2% to 29.2%), wage inequality will decrease by 0.28%. These results seem to show that to reduce wage inequality one should promote more equal access to training. In other words, policies aiming at increasing investment in training would not lead to a reduction in the dispersion of earnings if they do not change the process by which trainees are selected.

On the contrary, we already observed that discrimination against women is, ceteris paribus, a high factor of wage inequality between trainees and nontrainees. In fact, one can see that if discrimination against women disappeared, that is, if the proportion of women in the group of trainees

(43.4%) increases to correspond to the proportion of women in the sample (46.3%) (see Table 2), the proportion of trained women would increase by 2.9% and this would lead to a very important reduction (20%) in the between groups variance.

Similarly, if the proportion of individuals in the group of trainees (30.7%) working in firms having 10–49 workers increases and becomes equal to the share of these individuals in the whole sample (33.4%), the between-group variance would decrease by 20.3%. As large firms offer higher wages, because they already employ higher-skilled individuals (Abowd & Kramarz, 1998), this result shows that large firms tend to train more. This importance of firm size in the training selection process confirms the existence of segmented markets in France.

The same kind of remarks may be made for the educational level of the individuals or the level of qualification of their job. Thus if the proportion of "BEPC" or "employees" in the group of trainees (when compared to that in the group of nontrainees) increases by 1%, the between-group variance will decrease by 4.1% and 4.3%, respectively.

It should by now be clear that the process by which trainees are selected is a high factor of wage dispersion between the two groups, when we compare it to the net effect of training. Policy recommendations aiming at reducing the wage gap between trainees and nontrainees may therefore include the introduction of positive discrimination and the promotion of training investment in small firms. We should, however, not forget that most of the dispersion of earnings takes place within groups and may be explained by the heterogeneity of returns on training. We analyze this aspect in the following section.

5.2. Contributions of the Various Variables to the Within-Group Variance: The Case of Trainees

Table 5 gives the contribution of the various explanatory variables to the within-group dispersion in earnings. Here also, in the case of variables that have several modalities, we sum the contributions of all the modalities. Differences between the contributions observed for trainees and nontrainees confirm the heterogeneity of training programs as far as their returns or the selection process are concerned.

First of all, we note that the contribution of gender to the dispersion of earnings is smaller among trainees than among nontrainees. This implies either that that the returns to training are higher for women or that women are more likely to be selected in training programs that offer higher returns.

Table 5. Decomposition of the within Variance of the Logarithm of Wages.

Variable	Trainees		Nontrainees		Within Contribution (%)
	Contribution (%)	SE	Contribution (%)	SE	
Trimester	0.065	(0.190)	0.035	(0.107)	0.043
Number of children	0.009	(0.080)	−0.028	(0.041)	−0.018
Women	1.329	(0.417***)	1.567	(0.457***)	1.5
Foreigner	−0.004	(0.034)	−0.036	(0.050)	−0.027
Age and seniority	2.133	(0.507***)	1.390	(0.338***)	1.6
Education level	4.7	(0.992***)	2.971	(0.539***)	3.459
Working hours	12.349	(2.092***)	23.029	(2.046***)	20.015
Job Contract	0.323	(0.423)	0.164	(0.251)	0.209
Job schedule	−0.222	(0.259)	0.117	(0.116)	0.021
Qualification level	10.727	(1.440***)	5.536	(0.791***)	7
Position	0.677	(0.696)	1.593	(0.847**)	1.289
Sector	1.790	(0.557***)	1.297	(0.423***)	1.436
Firm size	1.864	(0.487***)	2.890	(0.657***)	2.6
Paris	0.618	(0.265***)	0.582	(0.240***)	0.592
New equipment	0.465	(0.296*)	0.279	(0.171*)	0.331
New working organization	−0.048	(0.129)	−0.042	(0.112)	−0.044
New job	−0.181	(0.172)	−0.089	(0.089)	−0.115
Dummies of position in wage distribution	42.183	(2.122***)	36.394	(1.842***)	38.028
1st deciles	19.427	(2.686***)	36.049	(2.047***)	31.358
1st deciles to 1st quartile	14.539	(1.969***)	9	(1.111***)	10.563
1st quartile to 2nd quartile	10.819	(1.579***)	1.337	(0.689**)	4.013
2nd quartile to 3rd quartile	2.534	(0.977***)	−5.844	(0.537***)	−3.480
3rd quartile to 9th deciles	−5.136	(0.713***)	−4.147	(0.498***)	−4.426
Residual	21.237	(1.88***)	22.350	(1.012***)	22.036

Note: Standard errors were derived from a bootstrap procedure with 1,000 replications. Significant differences within trainees and nontrainees:
*10% level of significance.
**5% level of significance.
***1% level of significance.

Similar conclusions may be drawn for the impact of the size of the firm. In other words, although larger firms usually propose a higher pay, training in smaller firms seems to give higher returns. This seem to imply that small firms offer less training than larger firm but their training programs are more efficient. The impact of working hours and of the position of the worker may be analyzed in a similar way.

As mentioned previously, the total contribution of gender and firm size to the dispersion of earnings is equal to the sum of their contribution to the between-group variance and of the weighted sum of their contribution to the variance within each of the two groups (see expression (17)). We then obtain a contribution of 1.06% for gender and of 2.76% for the size of the firm. These relatively small contributions, however, reflect two opposite effects. On one hand, investment in training increases the dispersion of earnings because the selection of trainees depends on the gender of the individual and the size of the firm in which he/she works. On the other hand, participation to training reduces the dispersion of earnings because of the heterogeneity in returns to training.

It should, however, be clear that the heterogeneity in returns to training may also contribute to an increase in the dispersion of earnings. This is in fact the impact of variables such as age, seniority, the level of education, the level of qualification of the job, the sector, the area where the firm is located, and the introduction of new equipment.

One may be surprised by the impacts of the levels of education and of qualification of the job since these variables increase inequality within the group of trainees. These results are different from those mentioned in the report of OECD (1999), which assumed that returns to training would be higher for the less-educated individuals. One should, however, remember that our focus is on general training and therefore one may indeed expect that in such a case more able workers receive higher returns on their training. This could explain why firms have incentives to train the most-qualified workers. It may also be the case that if the value of the output depends more on the human capital of the most-qualified workers, training schemes that would allow better managerial decisions may lead to higher returns on training for those who have the most-qualified jobs.

The contribution of the level of education and that of the level of qualification of the job explain, respectively, 4.7 and 10.7% of the variance of earnings within the group of trainees, the corresponding contributions for nontrainees being 3 and 5.5%, respectively. As a whole, education and the level of job qualification are important determinants of the overall wage dispersion (38% and 7%, respectively). A comparison of the contributions to the overall within-group variance of earnings with those to the variance of earnings among trainees illustrates well the impact of investment in training. These results clearly show the process by which trainees are selected, and the heterogeneity of returns to training increase by 0.8% the impact of education and by 1.5% the impact of the level of job qualification on the dispersion of earnings.

Also, note that the set of dummies measuring the positions of individuals in the wage distribution before the training takes place has a greater impact on the dispersion of earnings among trainees than nontrainees. We may therefore conclude that unobservable factors have still an important impact on the within-group dispersion in earnings. As mentioned previously, such an effect may well be due to the interaction of different characteristics.

Finally, we may note that the residuals have a significant effect on the within-group dispersion of earnings (approximately 22%). The contribution of this variable may reflect differences in the wage policy of firms. Firms may thus promote wage dispersion in order to create incentives and, as a consequence, induce a higher marginal productivity of workers (Lazear & Rosen, 1981). On the other hand, firms may limit wage inequality for fairness and equity reasons that may improve their performances (Akerlof & Yellen, 1990). As residuals have a positive effect on earnings dispersion, it is likely that the wage policy of firms leads to a higher dispersion of earnings.

6. CONCLUSIONS

The goal of this chapter was to estimate the exact impact of training on the dispersion of wages. We used an approach originally proposed by Fields (2003) but extended it to the breakdown of inequality by population subgroups. The empirical illustration was based on a survey conducted in France in 2004, 2005, and 2006.

The results of the analysis first show that when a distinction is made between workers who received training and those who did not, the between-group dispersion explains only 5.3% of the overall variance of earnings so that most of the dispersion in earnings turns out to be a within-group dispersion. It should therefore be clear, given that there is a small between-group dispersion and a big within-group dispersion, that there is a lot of overlapping between the distributions of earnings of the two groups, those who received and those who did not receive training. Such findings should imply that unobserved heterogeneity plays a key role in the selection of those who receive training and thus indirectly has an impact on the difference between the *average* earnings of those who receive and do not receive training. It cannot, however, be considered as a variable that could lie behind market segmentation. This is so because the within-group variance is much higher than that of between groups, so that the distribution of earnings of these two groups show a great degree of overlapping. In other words, there is a much greater degree of heterogeneity within than between

the two groups corresponding to those who received and did not receive on-the-job training. As a consequence if labor market segmentation exists, it must be based on other criteria.

Second, we have demonstrated that investment in general training affects earnings dispersion via three main channels. First, training has a small direct average impact on wage inequality since its contribution to the overall variance is of 0.7%. Second, training has a much stronger effect on the dispersion of earnings when the process by which trainees are selected is taken into account. It thus turns out that training raises the initial inequalities between genders, between educational and qualification levels, and between firms.

Third, investment in training can also have an impact on the dispersion of earnings via the heterogeneity of the returns on training. This effect depends, however, on the factors considered. As far as gender and the size of the firm are concerned, the heterogeneity of investment in training reduces inequality. It seems, however, that the inequality of wages between the different levels of education or socio-professional categories becomes higher rather than smaller after the training takes place.

We may therefore conclude that policies aiming at using vocational continuous training to reduce wage inequality should mainly focus on a better allocation of training expenses, one that would favor women, small firms, and the less-qualified workers, rather than try to increase the total amount of expenses on investment in training.

NOTES

1. For other studies stressing the unequal distribution of training, see, for example, Crocquey (1995), Aventur and Hanchane (1999), Blundell, Dearden, Meghir, and Sianesi (1999), and Ariga and Brunello (2003).

2. It is important to understand that the role of training may vary from one country to another. Thus, in Germany the educational system is such that the knowledge accumulated at school has a high productive value and there is thus little uncertainty about the skills of those who hold a diploma. Continuous training may then be considered as an additional way of improving the quality of the human capital of the workers and hence have a clear impact on earnings. In France, on the contrary, there is a lot of uncertainty about the skills of those who hold a diploma, especially at low and intermediate levels, so that firms will choose a strategy that progressively reveals the productive capacities of the workers. Such a matching process explains why access to training has to be selective and is mainly reserved to those workers who succeeded in overcoming the barriers to entry into internal markets.

3. For general studies of the causes of increasing wage dispersion, see, for example, Levy and Murnane (1992) or Karoly (1992). For studies emphasizing the

role of skill biased technological change, see, for example, Bound and Johnson (1992), Katz, Lawrence, and Murphy (1992), and more recently, Heckman and Lochner (1998) and Krusell, Ohanian, Rios-Rull, and Violante (2000).

4. Here evidently there is no contribution of factor F to the within groups variance.

5. See Wooldrige (2002) for a survey of average treatment effect methods.

6. Altonji and Spletzer (1991) and Harris (1999) present several determinants of training participation.

7. Similar findings about the effect of seniority in France may be found in the works of Béret (1992), Goux and Maurin (1994), and Hanchane and Joutard (1998). These results are an illustration of the transformations that occurred in the French labor market as well as of its specificity when compared with other industrial countries. Before what is known in France as the "crisis," which started in the mid 1970s, there was a close link between the worker and his job. Qualification was thus acquired progressively while working. The "crisis," which led to a stronger emphasis on competitiveness, put in evidence the rigidity of internal markets so that external markets became the preferred choice of those individuals who had acquired a minimal level of investment in education. As a consequence, though seniority increased, its return decreased, even sometimes becoming nil. Various studies such as those of Maurice, Sellier, and Silvestre (1982), Silvestre (1986), Verdier (1997), and Béret (1992) have actually emphasized these transformations of the French labor market.

REFERENCES

Abowd, J., & Kramarz, F. (1998). Internal and external labor markets: An analysis of matched longitudinal employer-employee data. In: J. Haltiwanger, M. Manser & T. Topel (Eds), *Labor statistics measurement issues*. University of Chicago Press.

Acemoglu, D. (1997). Training and innovation in an imperfect labour market. *Review of Economic Studies*, *64*, 445–464.

Acemoglu, D., & Pischke, J. (1998). Why do firms train? Theory and evidence. *Quarterly Journal of Economics*, *113*(1), 79–119.

Acemoglu, D., & Pischke, J. (1999a). Beyond Becker: Training in imperfect labour markets. *The Economic Journal*, *109*, 112–142.

Acemoglu, D., & Pischke, J. (1999b). The structure of wages and investment in general training. *Journal of Political Economy*, *107*(3), 539–572.

Acemoglu, D., & Pischke, J. (2003). Minimum wages and on-the-job training. *Research in Labor Economics*, *22*, 159–202.

Akerlof, G., & Yellen, J. L. (1990). The fair wage-effort hypothesis and unemployment. *Quarterly Journal of Economics*, *105*(2), 255–283.

Almeida-Santos, F., & Mumford, K. (2006). *Employee training, wage dispersion and equality in Britain*. Discussion paper no. 2006/14. Department of Economics, University of York, Heslington, York.

Altonji, J., & Spletzer, J. (1991). Worker characteristics, job characteristics, and the receipt of on-the-job training. *Industrial and Labor Relations Review*, *45*(1), 58–79.

Ariga, K., & Brunello, G. (2003). *Education, training and productivity: Evidence from Thailand and the Philippines*. Empirical Analysis of Economic Institutions Discussion Paper Series No. 13. Kyoto Univeristy.

Aventur, F., & Hanchane, S. (1999). Justice sociale et formation continue dans les entreprises françaises. *Formation Emploi*, *66*, 5–20.

Becker, G. (1964). *Human capital: A theoretical analysis, with special reference to education*. New York: Columbia University Press.

Béret, P. (1992). Salaires et marchés internes. *Économie Appliquée*, *XLV*(2), 5–22.

Béret, P., & Dupray, A. (2000). Allocation et effet salarial de la formation professionnelle continue en France et en Allemagne: une approche en terme d'information. *Economie Publique*, *5*, 221–269.

Blundell, R., Dearden, L., Meghir, C., & Sianesi, B. (1999). Human capital investment: The returns from education and training to the individual, the firm and the economy. *Fiscal Studies*, *20*(1), 1–23.

Booth, A., & Chatterji, M. (1998). Unions and efficient training. *The Economic Journal*, *108*, 328–343.

Bound, J., & Johnson, G. (1992). Changes in the structure of wages in the 1980s: An evaluation of alternative explanations. *American Economic Review*, *82*, 371–392.

Chang, C., & Wang, Y. (1996). Human capital investment under asymmetric information: The Pigovian conjecture revisited. *Journal of Labor Economics*, *14*, 505–519.

Crocquey, E. (1995). La formation professionnelle continue: des inégalités d'accès et des effets sur la carrière peu importants à court terme. *Travail et Emploi*, *65*, 61–68.

Fields, G. (2003). Accounting for income inequality and its change: A new method, with application to the distribution of earnings in the United States. *Research in Labor Economics*, *22*, 1–38.

Freeman, R. (1984). Longitudinal analyses of the effects of trade unions. *Journal of Labor Economics*, *2*, 1–26.

Goux, D., & Maurin, E. (1994). Education, expérience et salaire: tendances récentes et évolution de long terme. *Economie et Prévision*, *116*(5), 155–178.

Goux, D., & Maurin, E. (1997). Les entreprises, les salariés et la formation continue. *Economie et Statistique*, *306*, 41–55.

Hanchane, S., & Joutard, X. (1998). Une approche empirique de la structure du marché du travail: Salaires, formes de mobilité et formation professionnelle continue. *Economie et Prévision*, *135*, 57–75.

Harris, R. (1999). The determinants of work-related training in Britain in 1995 and the implications of employer size. *Applied Economics*, *31*, 451–463.

Heckman, J. (1979). Sample specification bias as a specification error. *Econometrica*, *47*(1), 153–161.

Heckman, J., & Lochner, L. (1998). Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics*, *1*, 1–58.

Karoly, L. A. (1992). Changes in the distribution of individual earnings in the United States: 1967–1986. *Review of Economics and Statistics*, *74*(1), 107–115.

Katz, E., Lawrence, F., & Murphy, K. (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics*, *107*(1), 35–78.

Katz, E., & Ziderman, A. (1990). Investment in general training: The role of information and labour mobility. *The Economic Journal*, *100*(403), 1147–1158.

Krusell, P., Ohanian, L., Rios-Rull, J.-V., & Violante, G. (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica*, *68*, 1029–1054.

Lazear, E. (2003). *Firm-specific human capital: A skill-weights approach*. Working Paper no. 9679 NBER, Cambridge, MA.
Lazear, E. P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, *89*(5), 841–864.
Levy, F., & Murnane, R. (1992). US Earnings levels and earnings inequality: A review of recent trends and proposed explanations. *Journal of Economic Literature*, *30*, 1333–1381.
Loewenstein, M., & Spletzer, J. (1998). General and specific training. *The Journal of Human Resources*, *34*(4), 710–733.
Maurice, M., Sellier, F., & Silvestre, J.-J. (1982). *Politique de l'éducation et organisation industrielle en France et en Allemagne. Essai d'Analyse sociétale*, PUF.
Mood, A. M., Graybill, F. A., & Boes, D. C. (1974). *Introduction to the theory of statistics* (3rd ed.). Auckland: McGraw-Hill.
OECD. (2005). *Promouvoir la formation des adultes*, Editions OCDE, Paris.
OECD. (1999). *Perspectives de l'emploi* – Chapter 3: Formation des travailleurs adultes dans les pays de l'OCDE: mesure et analyse.
Silvestre, J.-J. (1986). Marchés du travail et crise économique : de la mobilité à la flexibilité. *Formation Emploi*, *14*, 54–61.
Stevens, M. (1994). A theoretical model of on-the-job training with imperfect competition. *Oxford Economic Papers*, *46*(4), 537–562.
Verdier, E. (1997). Insertion des jeunes à la française: vers un ajustement structurel? *Travail et Emploi*, *69*, 37–59.
Wooldrige, J. (2002). *Econometric analysis of cross section and panel data*. New York: Columbia University Press.

APPENDIX A. ADDITIONAL DETAILS ON THE METHODOLOGY

The contribution of the explanatory variables to the variance of earnings may be analyzed as follows, according to Fields (2003).

Recalling that the earnings function is expressed as

$$y_i = \sum_{k=1}^{K+2} b_k Z_{k,i} \tag{A.1}$$

we derive

$$\mathrm{Var}(y_i) = \mathrm{Cov}\left(\sum_{k=1}^{K+2} b_k Z_{k,i}; y_i\right) \tag{A.2}$$

Dividing both sides of (A.2) by $\mathrm{Var}(y_i)$, we then derive that

$$1 = \frac{\mathrm{Cov}(\sum_{k=1}^{K+2} b_k Z_{k,i}; y_i)}{\mathrm{Var}(y_i)} \tag{A.3}$$

It is, however, well known (see Mood, Graybill, & Boes, 1974) that

$$\mathrm{Cov}\left(\sum_{k=1}^{K+2} b_k Z_{k,i}; y_i\right) = \sum_{k=1}^{K+2} \mathrm{Cov}(b_k Z_{k,i}; y_i) \tag{A.4}$$

Expression (A.3) may therefore be expressed as

$$1 = \frac{\sum_{k=1}^{K+2} \mathrm{Cov}(b_k Z_{k,i}; y_i)}{\mathrm{Var}(y_i)} \tag{A.5}$$

since $\mathrm{Cov}(\sum_{k=1}^{K+2} b_k Z_{k,i}; y_i) = \sum_{k=1}^{K+2} \mathrm{Cov}(b_k Z_{k,i}; y_i) = \mathrm{Var}(y_i)$

If we also remember that the correlation coefficient between $b_k Z_{k,i}$ and y_i may be expressed as

$$\mathrm{Cor}(b_k Z_{k,i}; y_i) = \frac{\mathrm{Cov}(b_k Z_{k,i}; y_i)}{(\sigma(b_k Z_{k,i}), \sigma(y_i))} \tag{A.6}$$

we end up, combining expressions (A.3)–(A.6), with

$$1 = \frac{\sum_{k=1}^{K+2} \mathrm{Cor}(b_k Z_{k,i}; y_i)\sigma(b_k Z_{k,i})}{\sigma(y_i)} 1 \tag{A.7}$$

However, since

$$\mathrm{Cor}(b_k Z_{k,i}; y_i) = \mathrm{Cor}(Z_{k,i}; y_i) \tag{A.8}$$

expression (A.7) implies that

$$\sigma(y_i) = \sum_{k=1}^{K+2} (b_k) \mathrm{Cor}(Z_{k,i}; y_i)\sigma(b_k Z_{k,i}) \tag{A.9}$$

The between-group variance V_{BET} may be expressed as

$$V_{\mathrm{BET}} = f(\bar{y}_A - \bar{y})^2 + (1-f)(\bar{y}_B - \bar{y})^2 \tag{A.10}$$

$$\Leftrightarrow V_{\mathrm{BET}} = f(\bar{y}_A)^2 + (1-f)(\bar{y}_B)^2 - (f\bar{y}_A + (1-f)\bar{y}_B)^2 \tag{A.11}$$

since $\bar{y} = f\bar{y}_A + (1-f)\bar{y}_B$. Expression (A.10) may then be easily simplified to finally derive that

$$V_{\mathrm{BET}} = f(1-f)(\bar{y}_A - \bar{y}_B)^2 \tag{A.12}$$

APPENDIX B. RESULTS OF THE ESTIMATION OF WAGE REGRESSIONS

Variable	Model 1		Model 2		Model 3		Model 4	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Intercept	6.931	(0.105***)	0.041	(0.011***)	0.135	(0.090)	7.665	(0.095***)
Training access	0.044	(0.013***)	0.027	(0.011**)	0.0361	(0.011***)	0.0376	(0.012***)
Trimester 2	−0.013	(0.016)	0.009	(0.014)	0.007	(0.014)	−0.007	(0.014)
Trimester 3	0.009	(0.016)	−0.003	(0.014)	−0.008	(0.014)	−0.003	(0.014)
Trimester 4	0.014	(0.016)	−0.01	(0.014)	−0.008	(0.014)	0.0003	(0.014)
Foreigner	0.021	(0.025)			−0.002	(0.021)	0.018	(0.022)
Women	−0.12	(0.014***)			0.017	(0.012)	−0.049	(0.013***)
Number of children	0.021	(0.011*)			0.013	(0.01)	0.018	(0.01*)
Square of number of children	−0.003	(0.003)			−0.005	(0.003*)	−0.004	(0.003)
Age	0.01	(0.005**)			−0.006	(0.004)	−0.0003	(0.004)
Square of age	−1E−4	(0.0001)			0.0001	(0.0001)	0.00003	(0.0001)
Seniority	0.001	(0.0002***)			0.00001	(0.0002)	0.001	(0.0002***)
Square of seniority	−1E−6	(4.2E−7***)			−9.9E−8	(3.6E−7)	−9.2E−7	(3.7E−7**)
Seniority unknown	0.116	(0.062*)			0.072	(0.053)	0.073	(0.054)
Seniority < 1.5 years	0.033	(0.027)	0.029	(0.019)	−0.002	(0.024)	0.03	(0.024)
BEPC	0.054	(0.015***)			−0.008	(0.013)	0.025	(0.013*)
Baccalaureate	0.132	(0.020***)			0.012	(0.018)	0.075	(0.018***)
General studies	0.022	(0.057)			0.017	(0.049)	0.016	(0.049)
Technological studies	0.195	(0.0252***)			−0.008	(0.022)	0.103	(0.022***)
Paramedical studies	0.328	(0.052***)			−0.013	(0.045)	0.16	(0.045***)
Grande école	0.285	(0.038***)			0.007	(0.033)	0.162	(0.033***)
Master and doctorate	0.287	(0.029***)			−0.009	(0.025)	0.161	(0.025***)

Construction	0.027	(0.025)			−0.008	(0.021)	0.015	(0.022)
Trade	−0.048	(0.021**)			0.018	(0.018)	−0.010	(0.018)
Education, health	−0.156	(0.025***)			−0.045	(0.022**)	−0.100	(0.022***)
Services	−0.008	(0.017)			−0.004	(0.015)	0.004	(0.015)
10–49 workers	−0.087	(0.019***)			0.017	(0.017)	−0.034	(0.017**)
50–199 workers	−0.069	(0.019***)			0.009	(0.017)	−0.028	(0.017*)
200–499 workers	−0.052	(0.021**)			−0.0002	(0.018)	−0.026	(0.019)
Firm size unknown	−0.225	(0.025***)			0.011	(0.021)	−0.124	(0.022***)
Paris	0.146	(0.025***)			0.018	(0.022)	0.086	(0.022***)
New equipment	0.024	(0.014*)	0.0004	(0.0113)	0.009	(0.012)	0.018	(0.012)
New organization	−0.005	(0.013)	−0.013	(0.0109)	−0.003	(0.011)	−0.004	(0.011)
Unemployment rate	−0.003	(0.003)			−0.002	(0.003)	−0.003	(0.003)
Work 35–39 hours	−0.12	(0.015***)			0.034	(0.013***)	−0.042	(0.013***)
Work 30–34 hours	−0.272	(0.026***)			0.017	(0.022)	−0.139	(0.023***)
Work 15–30 hours	−0.571	(0.024***)			−0.007	(0.021)	−0.312	(0.023***)
Less 15 hours	−1.379	(0.040***)			−0.003	(0.035)	−1.021	(0.038***)
Temporary contract	−0.056	(0.033*)			0.046	(0.028)	−0.024	(0.029)
Other contract	−0.099	(0.065)			0.007	(0.056)	−0.055	(0.057)
Flexible schedule	−0.005	(0.014)			−0.004	(0.012)	−0.003	(0.012)
Sunday work	0.036	(0.017**)			−0.010	(0.015)	0.014	(0.015)
Saturday work	0.007	(0.014)			−0.015	(0.012)	−0.004	(0.012)
Evening work	0.033	(0.018*)			−0.019	(0.016)	0.01	(0.016)
Night work	0.027	(0.019)			0.002	(0.016)	0.007	(0.017)
Home work	−0.058	(0.021***)			0.015	(0.018)	−0.015	(0.018)
Qualified worker	0.077	(0.020***)			0.009	(0.018)	0.037	(0.018**)
Employees	0.036	(0.022)			0.05	(0.019***)	0.036	(0.02*)
Technicians, supervisor	0.185	(0.024***)			0.015	(0.021)	0.080	(0.021***)
Engineer, executives	0.478	(0.029***)			0.017	(0.025)	0.225	(0.027***)
Director, Manager	0.621	(0.074***)			0.059	(0.064)	0.36	(0.066***)
Other qualification	0.082	(0.039**)			0.076	(0.034**)	0.067	(0.034**)

APPENDIX B. (*Continued*)

Variable	Model 1		Model 2		Model 3		Model 4	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Repairing, cleaning	0.020	(0.023)			0.019	(0.020)	0.02	(0.020)
Hygiene and security	−0.108	(0.03***)			0.028	(0.026)	−0.067	(0.026***)
Transport	−0.034	(0.024)			0.001	(0.020)	−0.023	(0.021)
Secretarial	0.057	(0.03*)			−0.074	(0.026***)	−0.015	(0.026)
Administration	0.107	(0.028***)			−0.055	(0.024**)	0.014	(0.025)
Trade	0.062	(0.026**)			−0.023	(0.022)	0.011	(0.022)
Research	0.060	(0.029**)			−0.009	(0.025)	0.012	(0.025)
Teaching	0.113	(0.032***)			0.013	(0.028)	0.062	(0.028**)
Position unknown	0.024	(0.025)			−0.008	(0.021)	0.004	(0.022)
Higher working hours	−0.033	(0.026)	0.18	(0.022***)	0.186	(0.022***)	0.072	(0.023***)
Smaller working hours	0.103	(0.028***)	−0.171	(0.023***)	−0.175	(0.024***)	0.006	(0.024)
Higher qualification	−0.024	(0.038)	0.060	(0.032*)	0.052	(0.033)	0.021	(0.033)
Smaller qualification	0.015	(0.046)	0.01	(0.039)	0.004	(0.04)	0.027	(0.040)
Positions different	−0.025	(0.026)	0.032	(0.021)	0.032	(0.022)	−0.005	(0.022)
Contract different	−0.005	(0.029)	−0.028	(0.025)	−0.028	(0.025)	−0.022	(0.025)
1st deciles							−0.973	(0.034***)
1st deciles–1st quartile							−0.66	(0.028***)
1st quartile–2nd quartile							−0.545	(0.025***)
2nd quartile–3rd quartile							−0.414	(0.023***)
3rd quartile–9th deciles							−0.221	(0.022***)
R^2	0.7235	0.0525	0.0776	0.7907				

Note: Significant differences between trainees and nontrainees:
*10% level of significance.
**at 5% level of significance.
***at 1% level of significance.

EMPLOYEE TRAINING AND WAGE DISPERSION: WHITE- AND BLUE-COLLAR WORKERS IN BRITAIN

Filipe Almeida-Santos, Yekaterina Chzhen and Karen Mumford

ABSTRACT

We use household panel data to explore the wage returns associated with training incidence and intensity (duration) for British employees. We find these returns differ depending on the nature of the training, who funds the training, the skill levels of the recipient (white- or blue-collar), the age of the employee and if the training is with the current employer or not. Using decomposition analysis, training is found to be positively associated with wage dispersion: a virtuous circle of wage gains and training exists in Britain but only for white-collar employees.

1. INTRODUCTION

Training is a key factor in the economic performance of all countries. It is a major tool for increasing productivity and living standards

Jobs, Training and Worker Well-Being
Research in Labor Economics, Volume 30, 35–60
Copyright © 2010 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 0147-9121/doi:10.1108/S0147-9121(2010)0000030005

(Ok & Tergeist, 2002). Concentrating training amongst workers who perform complex tasks and have high levels of formal education may create a virtuous circle for these high-skill workers resulting in higher wages, further training opportunities, longer tenure and greater social status (Gershuny, 2005). In contrast, workers who are disadvantaged in the education process may be less likely to receive training, inducing a vicious circle for these low-skill workers, further increasing their risk of unemployment and social exclusion (Keep, Mayhew, & Corney, 2002).

Simply ensuring equity of training opportunity may not be sufficient to assure a reduction in wage inequality amongst workers if individuals with different characteristics obtain different benefits from the same training scheme. The British government is increasingly concerned with the potentially contradictory implications of training policy for equity and efficiency, namely, redirecting training investment towards groups that typically receive less training or towards groups of workers where expected returns are larger (Department of Trade and Industry, 2005).

This chapter concentrates on the relationship between training and wages. We seek to address a fundamental question: what is the nature of the contribution of training (and its financing) to wage inequality in Britain? In the process of seeking answers to this question, it is important to estimate the individual employee's rates of wage return to training. Relevant empirical studies are not easy to locate; Frazis and Loewenstein (2005) recently concluded 'we are aware of few studies that attempt to estimate rates of return to training'. Often due to data constraints, most of the relevant studies that do exist estimate average returns for all training recipients, ignoring that the provision and returns to training across employees may differ according to gender, age, education level, occupation and sector of employment. Using longitudinal data on households and individuals (the British Household Panel Survey, BHPS), we can address many of these issues. Our results may also be seen as a further empirical investigation of the potential returns from training which help to fill a gap in a still unresolved area of research (Pischke, 2001, p. 543; Leuven, 2002, p. 34).

The modelling of wage returns to training is considered in Section 2 of the chapter, data and variable descriptions are discussed in Section 3, results are presented in Section 4, wage returns to training within skill groups are explored in Section 5 and Section 6 concludes the chapter.

2. MODELLING WAGE RETURNS

The relationship between investment in training and wages has been explored extensively by Becker (1962, 1964), Ben-Porath (1967) and, of

course, Mincer (1958, 1962, 1970, 1974) with the development of the well-known Mincer wage regression.

In subsequent years, authors have increased the number of explanatory variables included in the regression with the addition of variables capturing individual, job and firm characteristics (recent reviews are provided by Chiswick, 2003; Polachek, 2008). In this augmented framework, training may be considered as inherently heterogeneous and it is legitimate to expect the size of any associated wage returns to differ according to the nature and the type of the training programme (Leuven, 2004, p. 19). Several limitations have been identified in this research area associated with methodological questions, with database quality and with the mixed continuous-discrete nature of training variables. We will return to discuss some of these issues below.

Following the tradition of the literature on training (in particular, Loewenstein & Spletzer, 1998), we estimate the wage return from different types of training using the following Mincer-type wage regression:

$$\ln W_{ijt} = X_{ijt}\beta + Y_t\delta + T_{it-1}\alpha + \mu_i + v_{ij} + \varepsilon_{ijt} \quad (1)$$

where $\ln W_{ijt}$ is the natural logarithm of the real (2005 prices) hourly wage of individual i in job j at time t; $\mathbf{X}_{ijt}$ a vector of individual and job characteristics; T_{ij-1} represents single period lagged measures of training accumulated by the worker and $\mathbf{Y}_t$ is a vector of year-specific dummy variables. Unobserved characteristics are decomposed into an individual fixed effect (FE) μ_i, an unobserved job match specific component v_{ij} and a transitory shock ε_{ijt}. The individual effect μ_i is considered as an omitted measure of time invariant characteristics such as ability, motivation, and ambition or career commitment. The unobserved components (μ_i and v_{ij}) become a problem for the consistency of estimates if they are in some way correlated with the regressors. Following Loewenstein and Spletzer (1998), we address this problem by estimating the model with FEs and approximating v_{ij} with a binary variable accounting for employer change.

3. THE DATA

The data are taken from BHPS which is a nationally representative, annual sample of private British households. The BHPS was launched in 1991. Each year, individual adult members of households are interviewed over a broad range of socioeconomic topics resulting in a rich and relevant data set. In 1992 and 1993 respondents were asked for information on their lifetime employment status and job histories which are included in the analyses

below. The BHPS questionnaire was extended in (and continuously from) wave 8, conducted in 1998, to include information on the nature of the three most recent training courses attended since September of the previous year, and how these courses were financed. Our focus is on the 1998–2005 waves of data[1] as we are particularly interested in this training information, although information collected for individuals in all of the previous waves is included in the analysis presented below.

Our sample is an unbalanced panel of employed individuals in Britain, in the 18–65 age bracket (that are original, temporary or permanent BHPS sample members). We exclude those individuals whose relevant training information is missing, the minority of workers with no expected weekly working hours and those reporting working more than 75 h per week (including paid overtime). Any employed respondents with missing hourly earnings were excluded, as were those with missing data on any of the pertinent labour market or personal characteristics. Individuals with hourly earnings below £1 or exceeding £100 were also excluded from the analysis. Our final sample contains 34,900 training observations over eight years (1998–2005), from 8,862 individuals, a little over half of whom are women.

Concise variable definitions and summary statistics for the final sample are presented in Table 1. Means and standard deviations are presented in columns 1 and 2 for the full sample, and in columns 3 and 4 for those workers trained. Columns 5–8 (and columns 9–12) present analogous information for white-collar (and blue-collar) employees. We define the white-collar group of employees to be the: managerial, professional, associate professional and technical, sales and clerical and secretarial occupations. The blue-collar group consists of the: personal services, craft and related, plant and machine operatives and other semi-skilled and unskilled occupations.

3.1. Training Measures

The BHPS questionnaire asks individuals for information concerning the three most recent training events they have been on since September of the previous year. For each event they are asked if this training was:

1. To help you get started in your current job?
2. To increase your skills in your current job?
3. To improve your skills in your current job?
4. To prepare you for a job or jobs you might do in the future?
5. To develop your skills generally?

Table 1. Variable Definitions and Means (1999–2005).

	All				White-Collar				Blue-Collar			
	Mean	SD	With training		Mean	SD	With training		Mean	SD	With training	
			Mean	SD			Mean	SD			Mean	SD
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
(1) Individual employee characteristics												
Years of experience	15.20	10.98	14.30	10.35	14.88	10.60	14.14	10.15	15.74	11.58	14.67	10.79
Age	39.54	11.08	38.28	10.61	39.29	10.81	38.39	10.51	39.96	11.51	38.02	10.83
Married	0.59	0.49	0.57	0.49	0.59	0.49	0.57	0.50	0.60	0.49	0.59	0.49
Female	0.51	0.50	0.55	0.50	0.59	0.49	0.61	0.49	0.38	0.49	0.42	0.49
White	0.98	0.15	0.97	0.17	0.97	0.17	0.97	0.18	0.98	0.13	0.98	0.14
Having a child under 18 years	0.42	0.49	0.42	0.49	0.41	0.49	0.40	0.49	0.44	0.50	0.47	0.50
Years of school	10.80	3.14	11.55	2.88	11.62	2.95	12.13	2.74	9.41	2.97	10.26	2.75
Years of tenure	5.13	6.08	4.31	5.38	4.55	5.42	3.91	4.87	6.12	6.95	5.20	6.29
Log hours	3.49	0.38	3.52	0.34	3.49	0.34	3.51	0.31	3.50	0.43	3.52	0.39
Temporary job	0.03	0.17	0.03	0.18	0.03	0.18	0.03	0.18	0.03	0.17	0.03	0.17
Part time	0.18	0.39	0.15	0.36	0.18	0.38	0.15	0.35	0.19	0.39	0.17	0.37
Have a vocational qualification	0.42	0.49	0.46	0.50	0.44	0.50	0.46	0.50	0.40	0.49	0.46	0.50
Trained in previous 12 months	0.32	0.47	1.00	0.00	0.35	0.48	1.00	0.00	0.27	0.44	1.00	0.00
Number of training course – cumulated events 1998–2005	2.95	4.75	5.91	6.12	3.42	5.20	6.32	6.46	2.15	3.73	4.99	5.17
Participated in a general training course in the last year	0.28	0.45	0.87	0.34	0.31	0.46	0.87	0.33	0.23	0.42	0.85	0.35
Participated in a general training courses financed by employer in the last year	0.24	0.43	0.75	0.43	0.27	0.44	0.76	0.43	0.20	0.40	0.74	0.44
Number of general training course – cumulated events 1998–2005	1.91	2.68	3.81	3.16	2.18	2.83	4.00	3.24	1.45	2.33	3.36	2.93
With current employer	0.74	1.53	1.72	2.10	0.83	1.62	1.77	2.15	0.60	1.35	1.60	1.97
With previous employer	0.10	0.53	0.20	0.75	0.11	0.57	0.22	0.80	0.08	0.43	0.17	0.64

Table 1. (*Continued*)

	All				White-Collar				Blue-Collar			
	Mean	SD	With training		Mean	SD	With training		Mean	SD	With training	
			Mean	SD			Mean	SD			Mean	SD
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Number of general training course financed by the employer – cumulated events 1998–2005	1.65	2.50	3.31	3.07	1.89	2.65	3.49	3.16	1.26	2.16	2.93	2.82
With current employer	0.65	1.44	1.51	2.00	0.73	1.52	1.56	2.06	0.53	1.27	1.41	1.88
With previous employer	0.08	0.47	0.16	0.67	0.09	0.52	0.17	0.71	0.06	0.37	0.13	0.54
Days of training in previous 12 months	6.29	29.16	19.57	48.84	6.71	29.70	19.03	47.61	5.58	28.21	20.77	51.46
Days of training in a course with general components in the last year	5.53	27.56	17.19	46.49	5.86	27.94	16.60	45.09	4.97	26.90	18.49	49.44
Days of training in a course with general components financed by the employer in the last year	3.93	22.04	12.20	37.54	4.10	21.82	11.62	35.52	3.63	22.41	13.50	41.66
Union member	0.33	0.47	0.39	0.49	0.32	0.47	0.39	0.49	0.35	0.48	0.40	0.49
Changed employer in the last year – either for a better job or was dismissed	0.07	0.25	0.07	0.26	0.07	0.25	0.07	0.25	0.07	0.25	0.08	0.26
Promoted in the last year	0.06	0.25	0.10	0.29	0.08	0.28	0.11	0.32	0.03	0.18	0.05	0.23
Occupations												
Managers and administrators	0.15	0.35	0.15	0.36	0.23	0.42	0.22	0.42	0.00	0.00	0.00	0.00
Professionals	0.10	0.30	0.14	0.35	0.16	0.37	0.21	0.41	0.00	0.00	0.00	0.00
Associate professionals and technicians	0.13	0.33	0.18	0.38	0.20	0.40	0.26	0.44	0.00	0.00	0.00	0.00
Clerical and secretarial occupations	0.18	0.39	0.17	0.37	0.29	0.45	0.24	0.43	0.00	0.00	0.00	0.00
Craft and related occupations	0.10	0.30	0.08	0.26	0.00	0.00	0.00	0.00	0.26	0.44	0.24	0.43
Personal and protective services occupations	0.11	0.31	0.13	0.34	0.00	0.00	0.00	0.00	0.30	0.46	0.42	0.49
Sales and related occupations	0.07	0.26	0.05	0.21	0.11	0.32	0.07	0.25	0.00	0.00	0.00	0.00
Plants and machines operatives	0.09	0.28	0.06	0.23	0.00	0.00	0.00	0.00	0.24	0.42	0.19	0.39
Elementary occupations	0.07	0.26	0.05	0.21	0.00	0.00	0.00	0.00	0.20	0.40	0.15	0.36

(2) Workplace characteristics												
Economic sectors												
Agriculture, fishing, mining; electricity, gas and water	0.02	0.15	0.02	0.15	0.02	0.13	0.02	0.14	0.03	0.18	0.03	0.17
Manufacturing	0.18	0.39	0.14	0.35	0.13	0.33	0.11	0.31	0.28	0.45	0.22	0.41
Construction	0.04	0.20	0.04	0.19	0.02	0.15	0.02	0.15	0.07	0.26	0.07	0.26
Retail, wholesale, catering, hospitality	0.17	0.38	0.11	0.32	0.19	0.39	0.12	0.32	0.14	0.35	0.10	0.30
Transport, storage and communication	0.06	0.24	0.05	0.22	0.05	0.22	0.04	0.20	0.09	0.28	0.07	0.25
Financial intermediation, real state, renting and business activities	0.14	0.35	0.14	0.35	0.20	0.40	0.19	0.39	0.05	0.23	0.05	0.22
Public services and other sectors	0.37	0.48	0.49	0.50	0.40	0.49	0.50	0.50	0.34	0.47	0.46	0.50
Type of organizations												
Public organization	0.30	0.46	0.39	0.49	0.31	0.46	0.40	0.49	0.26	0.44	0.36	0.48
Private organization	0.67	0.47	0.56	0.50	0.64	0.48	0.54	0.50	0.71	0.45	0.61	0.49
Non-profitable organization	0.04	0.18	0.05	0.22	0.04	0.20	0.06	0.23	0.02	0.15	0.03	0.18
Region												
London	0.06	0.25	0.07	0.25	0.08	0.27	0.08	0.27	0.04	0.20	0.05	0.21
Size of workplace												
Fewer than 25 employees	0.33	0.47	0.29	0.45	0.32	0.46	0.28	0.45	0.35	0.48	0.31	0.46
25–49 employees	0.15	0.35	0.15	0.36	0.14	0.34	0.14	0.35	0.16	0.37	0.17	0.37
50–99 employees	0.12	0.32	0.12	0.33	0.11	0.32	0.12	0.32	0.13	0.33	0.12	0.33
100–199 employees	0.10	0.31	0.10	0.30	0.10	0.31	0.10	0.30	0.11	0.31	0.10	0.30
200–499 employees	0.13	0.34	0.13	0.34	0.13	0.34	0.13	0.34	0.13	0.33	0.13	0.34
500–999 employees	0.07	0.25	0.07	0.26	0.07	0.26	0.08	0.27	0.06	0.23	0.06	0.24
1,000+ employees	0.11	0.31	0.14	0.34	0.13	0.34	0.15	0.36	0.07	0.26	0.10	0.30
Real wage												
Real (2005 prices) wage	10.22	6.44	11.05	6.47	11.57	7.13	12.10	6.83	7.94	4.16	8.71	4.82
Log real (2005 prices) wage	2.18	0.52	2.27	0.50	2.30	0.53	2.37	0.50	1.97	0.43	2.06	0.45
Number of observations	34,900		11,221		21,937		7,740		12,963		3,481	

Based on the answers to this question, we define two categories of training for the construction of the dichotomous and continuous variables related to the incidence and intensity (duration) of training respectively. The first is the widest category including any of the five options and is defined simply as *training*. It consists of either specific and/or general training components, and is expected to improve the worker's skills either in their current job or in any other job.

The second category is defined as *general training*. In this category, the interviewees have chosen the fifth option and they explicitly recognized that the training event included a general component, however, this option choice is not mutually exclusive and they may have chosen other options too.

To construct the third and fourth measures, additional information concerning the financing of training is included. We define the third measure as *employer-financed training*, or simply *financed training*, and construct a binary indicator variable if the training event (options 1–5 above) was also financed by the employer. This variable is set equal to one if trained workers recognize that fees were paid by the employer or if they respond that there were 'no fees'. We similarly define the fourth measure as *employer-financed general training*, or simply *financed general training*, and construct a binary indicator variable that allows us to identify if the general training event (option 5 above) was also financed by the employer (or if the response was 'no fees').[2] In our sample, for more than 73% of courses attended in the current employer's workplace (or training centre) the workers involved reported no fees, implying the training was employer financed.

The proportion of employees responding that they had received training in Britain was just over 30% in the BHPS sample (column 1, panel 1 of Table 1). Amongst the specific group of trained individuals, 87% of the courses attended include components that explicitly improved their general skills whilst 75% of courses were additionally financed by the employer. On average, trained workers participated in 5.9 training courses over the eight years.

The average intensity (or duration) of the set of three training events attended per year was approximately 19.6 days. Not surprisingly, general training courses and financed general training events both tend to be of shorter duration. White-collar workers experience not only a higher training incidence but also an average higher intensity; this is true across age groups.

3.2. Individual and Job Characteristics

Amongst the group of variables quantifying individual and demographic characteristics, are several measures of the individual's aptitude and

opportunities which may be related to wages and training outcomes, such as labour market work experience, years of formal education, the possession of a vocational qualification, current job tenure, gender and race.

We use a continuous variable for the years of actual labour market work experience using the individual's employment history since first leaving full-time education.[3] This is a superior measure than the commonly used proxies of potential lifetime work experience (Polachek, 2008; Regan & Oaxaca, 2009). Table 1 also reveals that trained workers have more years of formal education and less years of tenure in their current job.[4]

It is important when investigating the relationship between training and wages to consider relationships that may otherwise limit the efficiency and/or consistency of training estimates. First, training accumulated in the current job should be distinctly measurable from training accumulated in previous jobs. This allows testing of the joint hypothesis of no depreciation and that training is transferable across employer. Furthermore, the measures of training incidence and intensity should ideally fully capture the amount of training accumulated over the working life because it is the stock of human capital accumulated via training, and not just by the most recent flow, that affects wages. This may be particularly pertinent for certain demographic groups, such as women. On average, and in contrast with the results obtained using British workplace data (Almeida-Santos & Mumford, 2005), women have a higher rate of participation in training programmes than men (35% and 30% respectively) in the BHPS sample.

We have data on the cumulated events of training acquired in the period 1998–2005. The stock of human capital accumulated before this period is captured by current job tenure and previous work experience at the beginning of the period. Using cumulated events allows for greater flexibility and reduces potential bias due to errors in self-reported training (Ariga & Brunello, 2006; Frazis & Loewenstein, 2005; Melero, 2004).

A further complication when calculating the return to training is related to promotions. It is possible that employees are offered training prior to being promoted and before increasing their job responsibilities; this potential correlation between job-related training receipt and future promotions also needs to be addressed (Melero, 2004, p. 14). The descriptive statistics in Table 1 indicate that individuals with longer working hours, current union membership, full-time employment status, vocational qualifications and who were promoted last year are more likely to be trained, especially in the case of women. We control for promotion in our estimations below.

Amongst the occupational groupings, managers and administrators; professional occupations and associated professional and technical

occupations are more likely to participate in a training programme compared to those employed in sales; plant and machines operators and elementary occupations. Suggesting that the likelihood to be trained may also increase with the task's complexity and the responsibility required for the job. To further explore this possibility, as discussed above, the sample is divided into white- and blue-collar workers.

It is assumed that white-collar workers are allocated to occupations where tasks are more complex and job responsibilities higher. White-collar workers usually enjoy faster wage growth, they are better educated, more able to perform intellectually complex work related tasks (Bishop, 1997) and consequently are predicted to generate a higher rate of return from training.

3.3. Workplace and Market Characteristics

Whilst non-work attributes may have a significant impact on training and productivity, the work environment characteristics beyond the control of employees may also inhibit ability and motivation to perform activities (Clifton, 1997). Several measures are included in the empirical analyses as controls for some of these characteristics: region, industrial sector, firm type (non-profit, privately owned) and firm size. The definitions and summary statistics for these workplace and market characteristics are included in the lower panels (panels 2 and 3) of Table 1.

4. RESULTS

Results for the estimates from the FE models for training incidence and intensity are presented in Tables 2–5. Though only the relevant wage returns are reported in these tables, the independent variables include the individual-level control variables listed in Table 1 and discussed in Section 3,[5] plus the more aggregate level controls (including the workplace characteristics[6] and year-specific dummy variables). A full list of the controls is provided in the endnotes to the tables and full estimation results are available from the authors upon request. All of the results are based upon robust standard errors.[7] Overall, the parameter estimates are generally well defined and have the expected sign.

Several alternative functional forms were also considered, with training measures entering quadratically, as a logarithm, a cubic root and incorporating interaction terms. However, neither robust results[8] nor higher

Table 2. Earnings and Training Incidence (FE).

Dependent Variable: Log of Real Hourly Wage	All		White-Collar		Blue-Collar	
Training incidence	(1)	(2)	(3)	(4)	(5)	(6)
Training_{t-1}	0.0063***	–	0.0076***	–	0.0010	–
	(0.0012)	–	(0.0014)	–	(0.0024)	–
Training_{t-1} in the current employer	–	0.0077***	–	0.0081***	–	0.0031
	–	(0.0017)	–	(0.0020)	–	(0.0034)
Training_{t-1} in the previous employer	–	−0.0016	–	0.0016	–	−0.0099
	–	(0.0054)	–	(0.0060)	–	(0.0122)
Employer financed training_{t-1}	0.0061***		0.0074***		0.0010	
	(0.0012)		(0.0014)		(0.0025)	
Employer financed training_{t-1} in the current employer	–	0.0081***	–	0.0085***	–	0.0033
	–	(0.0017)	–	(0.0020)	–	(0.0035)
Employer financed training_{t-1} in the previous employer	–	−0.0031	–	0.0008	–	−0.0117
	–	(0.0059)	–	(0.0062)	–	(0.0134)
General training_{t-1}	0.0074***	–	0.0088***	–	0.0016	–
	(0.0013)	–	(0.0015)	–	(0.0028)	–
General training_{t-1} in the current employer	–	0.0082***	–	0.0094***	–	0.0022
	–	(0.0019)	–	(0.0022)	–	(0.0039)
General training_{t-1} in the previous employer	–	0.0009	–	0.0022	–	−0.0077
	–	(0.0059)	–	(0.0066)	–	(0.0139)
Employer financed general training_{t-1}	0.0071***	–	0.0084***	–	0.0012	–
	(0.0013)	–	(0.0015)	–	(0.0026)	–
Employer financed general training_{t-1} in the current employer	–	0.0085***	–	0.0095***	–	0.0023
	–	(0.0019)	–	(0.0022)	–	(0.0041)
Employer financed general training_{t-1} in the previous employer	–	−0.0010	–	0.0017	–	−0.0155
	–	(0.0065)	–	(0.0071)	–	(0.0155)
Observations	34,900		21,937		12,963	

Source: British Household Panel Survey, 1998–2005.

Notes: Each entry in columns (1)–(6) measures marginal effects.

**Statistically significant at 90%.

***Statistically significant at 95% and above.

All of the results are based upon robust standard errors. Controls are also included for experience and experience squared, age and age squared, marital status, gender, race, having children, years of school, current job tenure and tenure squared, having permanent job, having a part time job, having vocational qualifications, being a union member, having changed employer in the previous 12 months, having been promoted with the same employer, year, economic sector, industry, size of workplace and region.

Table 3. Earnings and Training Intensity (FE).

Dependent Variable: Log of Real Hourly Wage	All		White-Collar		Blue-Collar	
Training intensity/100	(1)	(2)	(3)	(4)	(5)	(6)
$Training_{t-1}$	0.0249***	–	0.0275***	–	0.0161	–
	(0.0051)	–	(0.0062)	–	(0.0102)	–
$Training_{t-1}$ in the current employer	–	0.0284***	–	0.0321***	–	0.0195**
	–	(0.0064)	–	(0.0078)	–	(0.0117)
$Training_{t-1}$ in the previous employer	–	0.0087	–	0.0040	–	0.0092
	–	(0.0171)	–	(0.0227)	–	(0.0247)
Employer financed $training_{t-1}$	0.0308***	–	0.0343***	–	0.0173	–
	(0.0058)	–	(0.0063)	–	(0.0123)	–
Employer financed $training_{t-1}$ in the current employer	–	0.0361***	–	0.0396***	–	0.0223
	–	(0.0075)	–	(0.0089)	–	(0.0146)
Employer financed $training_{t-1}$ in the previous employer	–	0.0169	–	0.0157	–	0.0044
	–	(0.0198)	–	(0.0238)	–	(0.0419)
General $training_{t-1}$	0.0273***	–	0.0282***	–	0.0231***	–
	(0.0055)	–	(0.0067)	–	(0.0109)	–
General $training_{t-1}$ in the current employer	–	0.0301***	–	0.0325***	–	0.0233**
	–	(0.0070)	–	(0.0087)	–	(0.0130)
General $training_{t-1}$ in the previous employer	–	0.0075	–	−0.0051	–	0.0179
	–	(0.0178)	–	(0.0216)	–	(0.0268)
Employer financed general $training_{t-1}$	0.0329***	–	0.0337***	–	0.0247**	–
	(0.0059)	–	(0.0065)	–	(0.0129)	–
Employer financed general $training_{t-1}$ in the current employer	–	0.0373***	–	0.0392***	–	0.0270**
	–	(0.0080)	–	(0.0096)	–	(0.0157)
Employer financed general $training_{t-1}$ in the previous employer	–	0.0109	–	0.0022	–	−0.0026
	–	(0.0196)	–	(0.0212)	–	(0.0545)
Observations	34,900		21,937		12,963	

Source: British Household Panel Survey, 1998–2005.

Notes: Each entry in columns (1)–(6) measures marginal effects.

**Statistically significant at 90%.

***Statistically significant at 95% and above.

All of the results are based upon robust standard errors. Controls are also included for experience and experience squared, age and age squared, marital status, gender, race, having children, years of school, current job tenure and tenure squared, having permanent job, having a part time job, having vocational qualifications, being a union member, having changed employer in the previous 12 months, having been promoted with the same employer, year, economic sector, industry, size of workplace and region.

Table 4. Earnings and Training Incidence, by Age Group (FE).

Dependent Variable: Log of Real Hourly Wage	White-Collar			Blue-Collar		
Training incidence	<30	30–45	>45	<30	30–45	>45
	(1)	(2)	(3)	(4)	(5)	(6)
Training_{t-1} in the current employer	0.0073	0.0107***	0.0072***	0.0084	0.0080**	−0.0003
	(0.0057)	(0.0030)	(0.0034)	(0.0095)	(0.0047)	(0.0070)
Training_{t-1} in the previous employer	0.0091	−0.0000	−0.0104	0.0106	−0.0222	−0.0369
	(0.0119)	(0.0082)	(0.0175)	(0.0182)	(0.0250)	(0.0277)
Employer financed training_{t-1} in the current employer	0.0041	0.0110***	0.0073***	0.0039	0.0081**	0.0009
	(0.0061)	(0.0030)	(0.0033)	(0.0100)	(0.0046)	(0.0077)
Employer financed training_{t-1} in the previous employer	0.0075	−0.0004	−0.0164	0.0059	−0.0319	−0.0284
	(0.0117)	(0.0089)	(0.0175)	(0.0229)	(0.0272)	(0.0275)
General training_{t-1} in the current employer	0.0094	0.0124***	0.0072**	0.0060	0.0083	−0.0010
	(0.0057)	(0.0032)	(0.0039)	(0.0105)	(0.0051)	(0.0088)
General training_{t-1} in the previous employer	0.0084	0.0044	−0.0272	0.0044	−0.0149	−0.0454
	(0.0141)	(0.0089)	(0.0203)	(0.0196)	(0.0288)	(0.0341)
Employer financed general training_{t-1} in the current employer	0.0071	0.0119***	0.0082***	0.0025	0.0078	0.0011
	(0.0062)	(0.0033)	(0.0039)	(0.0112)	(0.0051)	(0.0098)
Employer financed general training_{t-1} in the previous employer	0.0063	0.0059	−0.0387**	−0.0032	−0.0405	−0.0339
	(0.0136)	(0.0096)	(0.0212)	(0.0245)	(0.0320)	(0.0352)
Observations	4,806	10,526	6,605	2,848	5,769	4,346

Source: British Household Panel Survey, 1998–2005.
Notes: Each entry in columns (1)–(8) measures marginal effects.
**Statistically significant at 90%.
***Statistically significant at 95% and above.
All of the results are based upon robust standard errors. Controls are also included for experience and experience squared, age and age squared, marital status, gender, race, having children, years of school, current job tenure and tenure squared, having permanent job, having a part time job, having vocational qualifications, being a union member, having changed employer in the previous 12 months, having been promoted with the same employer, year, economic sector, industry, size of workplace and region.

Table 5. Earnings and Training Intensity, by Age Group (FE).

Dependent Variable: Log of Real Hourly Wage	White-Collar			Blue-Collar		
Training days/100	<30	30–45	>45	<30	30–45	>45
	(1)	(2)	(3)	(4)	(5)	(6)
$Training_{t-1}$ in the current employer	0.0494***	0.0243**	0.0238**	0.0582***	0.0331***	−0.0247
	(0.0141)	(0.0130)	(0.0126)	(0.0233)	(0.0166)	(0.0220)
$Training_{t-1}$ in the previous employer	0.0737***	−0.0257*	0.0141	0.0253	−0.0005	−0.1619***
	(0.0355)	(0.0171)	(0.0774)	(0.0421)	(0.1028)	(0.0692)
Employer financed $training_{t-1}$ in the current employer	0.0450***	0.0264**	0.0429***	0.0728***	0.0390***	−0.0293
	(0.0160)	(0.0159)	(0.0108)	(0.0261)	(0.0185)	(0.0273)
Employer financed $training_{t-1}$ in the previous employer	0.0722***	−0.0330**	−0.0586	0.0011	0.0024	−0.4915***
	(0.0366)	(0.0191)	(0.0677)	(0.0994)	(0.0755)	(0.1998)
General $training_{t-1}$ in the current employer	0.0518***	0.0244**	0.0259**	0.0690***	0.0359***	−0.0141
	(0.0158)	(0.0145)	(0.0142)	(0.0251)	(0.0177)	(0.0250)
General $training_{t-1}$ in the previous employer	0.0408	−0.0238	0.0555	0.0315	0.0532	−0.1559***
	(0.0412)	(0.0176)	(0.0787)	(0.0546)	(0.0797)	(0.0651)
Employer financed general $training_{t-1}$ in the current employer	0.0485***	0.0082	0.0431***	0.0724***	0.0333	−0.0135
	(0.0177)	(0.0216)	(0.0109)	(0.0299)	(0.0228)	(0.0344)
Employer financed general $training_{t-1}$ in the previous employer	0.0370	−0.0291***	−0.0267	0.0063	−0.0876	−0.4549***
	(0.0432)	(0.0092)	(0.1273)	(0.1562)	(0.2011)	(0.1747)
Observations	4,806	10,526	6,605	2,848	5,769	4,346

Source: British Household Panel Survey, 1998–2005.
Notes: Each entry in columns (1)–(8) measures marginal effects.
**Statistically significant at 90%.
***Statistically significant at 95% and above.
All of the results are based upon robust standard errors. Controls are also included for experience and experience squared, age and age squared, marital status, gender, race, having children, years of school, current job tenure and tenure squared, having permanent job, having a part time job, having vocational qualifications, being a union member, having changed employer in the previous 12 months, having been promoted with the same employer, year, economic sector, industry, size of workplace and region.

goodness of fit measures were obtained compared to the results reported in Tables 2–5. (These additional results are available from the authors on request.)

4.1. Training Incidence

As discussed above, the relationship between training incidence and wages may vary across types of employees. To consider this possibility more fully, FEs wage regressions are estimated for the full sample of employees (columns 1 and 2 of Table 2) and for two separate worker groups: white-collar (columns 3 and 4) and blue-collar (columns 5 and 6). Columns 1, 3 and 5 present the 'base' results for lagged training incidence. In columns 2, 4 and 6 cumulated lagged training measures are split into training with current employer and training with previous employers.

Beginning with the results for the full sample of employees (Table 2, column 1), the incidence of a training course (ignoring the components that the course may include) is associated with a modest but significant increase of 0.63% in wages (column 1, panel 1); 0.61% if the training course is financed by the employer (panel 2).[9] The wage return to training that explicitly included a general component is associated with an increase of 0.74% (panel 3); 0.71% if the general training course is financed by the employer (panel 4).

Similar estimates of wage returns from training have been obtained by Lynch (1992a, 1992b) using the American National Longitudinal Survey of Youth Cohorts and by Schøne (2004) using the Norwegian Survey of Organisations and Employees. Arulampalam, Booth, and Bryan (2004), using the European Community Household Panel Series (which incorporates data from the BHPS for Britain), conclude that 'Britain, Denmark and Finland – are also amongst the countries with the lowest returns, of approximately one percent per event'.

Our estimated wage returns to training are, however, relatively low compared to those obtained by Booth and Bryan (2007) and Melero (2004) using the BHPS for the period of 1998–2000 and 1991–2002, respectively. There are some important differences between our approach and these earlier studies that may help to explain our lower estimates. In particular, we consider employees aged 18–65 (they included 16–65 year olds); we include public sector employees;[10] our sample period is substantially longer; we control for a larger set of independent variables; and, perhaps most importantly, we use broader definitions of training.[11]

Dividing training events into those with the current or previous employer (column 2 of Table 2) reveals that training events with the previous employer do not have a statistically significant relationship with current wages in the full sample estimates. Further dividing training with previous employer into (i) firm-financed training, (ii) general training and (iii) firm-financed general training (reading down column 2) confirms this result; we consistently find that training events with previous employers do not have a significant systematic relationship with wages for the full sample of British employees. In contrast, across all the training categories considered, training with current employer is associated with a modest but significant increase in wages for the full sample of British employees. This finding is consistent with the human capital model if, for example, skills received from training have depreciated and/or the skills acquired from training are not transferable across employers. We further explore the implications of these findings by considering the white- and blue-collar employees separately.

4.2. White- and Blue-Collar Employees

Beginning with the results for white-collar employees (columns 3 and 4 of Table 2), the wage returns associated with training incidence are similar in size to those found for the full sample, and training events with previous employers are again not found to be significantly associated with wages (column 4).

For the blue-collar sample (columns 5 and 6), the wage returns related to training incidence for all four categories are substantially lower than those found for the white-collar workers, however, they are also imprecisely estimated.[12] Indeed, for blue-collar workers, training events are not found to have a significant association with wages for any of the four training incidence measures considered. This is true for training events with the previous or current employer.[13]

To reiterate, the results in Table 2 for training incidence indicate three major findings: (1) for the full sample of British employees, the wage returns associated with a training event are small and positive, (2) training events with previous employers are not associated with wage gains and (3) blue-collar employees do not experience wage rises related to training events.

4.3. Training Intensity

The estimates of the FE models for training intensity (duration) are reported in Table 3, the results presented in the table are scaled by 100 and should be

interpreted accordingly. The results for the full sample of British employees (columns 1 and 2) are consistent with those found for training incidence. All four of the training measures are associated with wage increases (column 1). Furthermore, it is training with the current employer that is associated with wage growth (column 2). There is no significant evidence that training intensity with previous employers is related to wage rates for the full sample of British employees.

Dividing the workers into white- and blue-collar, the results again reveal that training is consistently and significantly positively related to wage changes for white-collar employees (columns 3–4). For these employees, the cumulated days of training (training intensity) with the current employer has a significant and positive relationship with wages (0.03% in column 4, allowing for the scaling). A white-collar employee undergoing a training programme (which includes general components) lasting for 20 days, with their current employee, may expect a wage increase of 0.6%, ceteris paribus. Training with previous employers is again found to have an insignificant association with wage, in contrast to cumulated training days with the current employer. For blue-collar workers (columns 5 and 6), there is evidence that training intensity (duration) is associated with higher wage returns especially for training that includes an explicit general component. Wage returns from training intensity are, however, typically small and less well defined for blue-collar workers in contrast to those found for white-collar workers.

In summary, our results indicate that whilst wage returns from training events (incidence and intensity) with the current employee are consistent and significantly positive for white-collar employees, this is not the case for blue-collar employees. When positive and significant relationships are found for blue-collar workers, the wage returns are consistently low for these employees. Equal access to training programmes will not reverse wage inequality in favour of low-skilled employees if blue-collar employees do not derive a wage benefit from participating in training.

4.4. Decomposing the Wage Differential

It appears that training may have a non-negligible role in wage inequality amongst workers in Britain. We next evaluate the contribution of different types of training to wage dispersion during the time period. Following Oaxaca and Ransom (1994), the mean wage gap can be

written as:

$$\ln W_{\mathrm{w}} - \ln W_{\mathrm{b}} = \underbrace{\underbrace{(\bar{X}_{\mathrm{w}} - \bar{X}_{\mathrm{b}})\hat{\beta}^{*}}_{\ln(Q_{\mathrm{wb}}+1)}}_{\textit{Explained}\ \text{Part}\ (E)} + \underbrace{\underbrace{\bar{X}_{\mathrm{b}}(\hat{\beta}^{*} - \hat{\beta}_{\mathrm{b}})}_{\ln(\hat{\partial}_{\mathrm{w}*}+1)} + \underbrace{\bar{X}_{\mathrm{w}}(\hat{\beta}_{\mathrm{w}} - \hat{\beta}^{*})}_{\ln(\hat{\partial}_{\mathrm{b}*}+1)}}_{\textit{Unexplained}\ \text{Part}\ (U)} \qquad (2)$$

where W_{w} represents the wages of the white-collar group (advantaged group) and W_{b} the wages of the blue-collar group (disadvantaged group); $\ln(Q_{\mathrm{wb}}+1)$ is the endowment component; $\ln(D_{\mathrm{wb}}+1) = \ln(\delta_{\mathrm{w}*}+1) + \ln(\delta_{\mathrm{b}*}+1)$ the remuneration or the discrimination component; $\delta_{\mathrm{w}*}$ and $\delta_{\mathrm{b}*}$ are respectively the blue-collar wage disadvantage and white-collar wage advantage associated with discrimination and β^{*} is a set of benchmark coefficients equal to:

$$\beta^{*} = \Omega\hat{\beta}_{\mathrm{w}} + (I - \Omega)\hat{\beta}_{\mathrm{b}} \qquad (3)$$

representing a matrix of relative weights of the estimated vector of coefficients and the identity matrix (**I**). A range of other choices have been suggested for the weighting matrix Ω (Reimers, 1983; Cotton, 1988; Neumark, 1988; Oaxaca & Ransom, 1994).

The gross wage differential between white- and blue-collar workers across the time period is 33 log wage points. We find training is associated positively with wage dispersion[14] and that the type of training is itself of little relevance for wage dispersion: our widest category of training contributes little more than 1.9% of the overall wage differentials. Cumulated training events that explicitly include general training, either financed by the employer or not, reveal a higher but still modest contribution (of up to 2.9%).

The results do not suggest that training is a major tool for reversing wage inequality amongst workers. On the contrary, it seems that training is a contributor to the wage dispersion across white- and blue-collar workers, even if the training programme explicitly includes general components that may be expected to increase the employee's wage offers across firms.[15] The implications of these findings may be further explored by concentrating analyses on the returns to training for workers within skill and age bands.

Other studies have also found training to be positively associated with inequality, beginning with the seminal work on income inequality and training by Chiswick and Mincer (1972). Proxying training with years of labour market experience, they find cyclical changes in employment had a major impact on labour market experience and on income inequality in the

United States from 1939 (particularly so during years of the Great Depression). Whilst they had limited data, they found income inequality was sensitive to the rate of return to human capital (Chiswick & Mincer, 1972, p. S47) and to the age distribution of employees. They concluded that it is important to concentrate analyses on the returns to training for workers within skill and age bands.

The implications of the human capital model for the relationship between trainings and earnings over the life cycle (as the employee ages) are well known (Ben-Porath, 1967): the time spent in post-school investment in human capital declines monotonically over the life cycle, implying that the individual's stock of human capital is increasing but at a decreasing rate and that earnings will increase with age but at a decreasing rate. Introducing depreciation into the model implies that skills may eventually depreciate at a faster rate than investment and that net human capital will decline. Earnings functions are predicted to be concave with respect to age, and the logarithmic earnings profile to be U shaped (Polachek, 2003).

Periods of unemployment may constrain the process of post-school investment in human capital, especially if this investment is primarily associated with on the job training, resulting in flatter earnings functions over the life cycle (Polachek, 1975). These periods on unemployment may be largely expected (e.g. by those who plan to take time off to raise children or those working in sectors that are particularly affected by cyclical downturns or seasonal demand fluctuations) or they may be largely unpredicted (due, e.g. to structural change in the economy or periods of poor health). As the individuals age, the accumulated effects of such spells become more pronounced and greater earnings differences are found (Chiswick & Mincer, 1972).

The process may be confounded with the introduction of occupation-based differences in skill acquisition. For example, those expecting to spend time out of the labour market to rear children may choose occupations with a lower skill depreciation rate and/or skills that are also associated with increased productivity in their expected non-market activity (such as the nurturing skills associated with infant year school teachers). If structural changes in the economy (such as skilled bias technological changes) have varying affects on different occupations intertemporally a more complex relationship between earnings, training, age and occupation may arise. The inclusion of a more extensive range of control variables in the augmented model (such as actual work experience, gender, occupation and parental status) help to alleviate potential omitted variable bias but would not be expected to fully capture these confounding effects. In an attempt to further

address some of these issues, we follow the example of Chiswick and Mincer (1972) and next concentrate analyses on the returns to training for workers within skill and age bands.

5. WAGE RETURNS TO TRAINING WITHIN GROUPS

The white- and blue-collar groups are further subdivided into three different age groups: younger than 30, between 30 and 45 and older than 45 years (i.e. <30, 30–45 and >45). Table 4 presents the estimated wage returns from cumulated training incidence split into training with current employer and training with previous employers for the white-collar age groups (columns 1–3) and for the blue-collar age groups (columns 4–6). The models presented in Table 4 are directly comparable to those in Table 2, with the same set of control variables included (as listed in the endnotes of the tables).

A striking result is found when different age bands of white- and blue-collar employees are examined. Cumulated training events are not statistically significantly related to wages for either white- or blue-collar workers who are younger than 30years. As discussed above, the human capital model would predict that if an investment in human capital via training is profitable, it will have a higher net value the earlier it is undertaken, thereafter declining over time (Chiswick & Mincer, 1972, p. S37). The opportunity cost of engaging in training is also typically lower for younger employees whose intertemporal earnings functions are generally still rising (Polachek, 2008, p. 174). The result that training events are not significantly related to wages for young employees in Britain is therefore surprising.

Considering the white-collar employees in more detail, cumulated training events with the *current* employer are found to be significantly related to wage increases for these employees who are aged over 30. Whilst, all four of the training categories with the *previous* employer have a negative, but rarely significant, relationship with wages with white-collar workers who are aged over 45 (as does total cumulated training with the previous employee for those aged over 30).

For blue-collar employees, only training events with current employers are associated with wage growth and this is true only for those employees aged between 30 and 45 (at a significance level of 10%). Training events which explicitly include general training have a similar relationship with wages for these workers but with less precision (at a significance level of 15%). All four of the training categories with the *previous* employer have a negative, but insignificant, relationship with wages for blue-collar workers

aged over 30. No significant relationship is found between training incidence and wages for the younger or older groups of blue-collar workers.

Turning to consider training intensity (duration) by age and skill group, these results (scaled by 100) are presented in Table 5 (and are comparable to the results in Table 3). We again find that training intensity with current employee is positively associated with wages for white-collar employees (columns 1–3). For these workers who are aged over 30, training with previous employer is typically associated with negative but insignificant returns. For the younger white-collar workers (aged under 30), training intensity with previous employee is related to higher wages but only significantly so for the measures of total accumulated training duration (with and without firm financing).

The results for training intensity are more extreme for blue-collar workers (columns 4–6). For older blue-collar workers there is no significant gain associated with training intensity with their current employer and there are significant negative returns for training intensity with their previous employer (regardless of training type or financing). For younger blue-collar workers there are only significant returns associated with training intensity with current employer.

Taken in combination with the results presented in Table 4, for younger employees (white- or blue-collar) it is important that their training events are more intense (longer duration) and that this training takes place with their current employer. For those aged between 30 and 45, longer training with current employee is generally associated with higher returns for both skill groups. Whilst this is also true for older white-collar employees, we find no return from training (regardless of duration) for older blue-collar employees. Indeed, blue-collar employees are found to face a significant wage penalty associated with more intense training having occurred with previous employees.

The results reveal that the relationship between training (incidence and intensity), is not uniform for white- and blue-collar employees nor is it constant over the working life of an employee. Consequently, the impact of training policy may be distinct and/or have very different impacts with respect to the age and the occupation of the recipients.

6. CONCLUSION

We use British household panel data from 1991 to 2005 to explore the wage returns associated with training (both incidence and intensity) undertaken

by employees between 1998 and 2005. We find (after controlling for a range of individual and workplace characteristics) estimated wage returns to a training incident for British employees are typically small at less than 1%. Although, training courses that include general components are associated with a higher wage as are training courses undertaken with the current employer.

The relationship between training and wages is also found to differ according to the occupation (white- or blue-collar) and the age of the group of workers that participates in the training programme. We find very limited evidence of wage returns from training incidence for blue-collar employees. This result contrasts with the range of positive returns found for older (aged over 30) high-skill employees. Training intensity (duration) with current employer is found to be important for all white-collar and for younger (aged below 45) blue-collar employees. However, we find no evidence that training (incidence or duration) is associated with higher wages for older (above 45 years) blue-collar employees.

White-collar employees are shown to have higher training incidence and intensity than do blue-collar workers, suggesting a virtuous circle between training and wage growth for white-collar employees (but not for blue-collar employees). Using decomposition analysis, unequal returns associated with training for different skill groups are found to contribute modestly to wage inequality across white- and blue-collar employees in Britain. These results imply that promoting equal access to training programmes will not reverse wage inequality in favour of blue-collar workers. Indeed, it may exacerbate wage inequality.

NOTES

1. The latest wave of the BHPS data (2006/2007) was released in late September 2008, however, the introduction of new definitions and coding in this latest wave limited our ability to use this wave of data at the time of carrying out the analysis for this chapter.

2. Recent non-competitive models emphasize how market frictions may transform what the human capital model classifies to be general training into de facto specific training (Acemoglu & Pischke, 1999). In such an environment, firms have an incentive to finance general training and to distribute these training opportunities amongst employees, thereby introducing issues of allocation.

3. More specifically, we used the *BHPS Combined Work-Life History Data 1990–2005* (see Halpin, 2006). This dataset combines information about the current activity status of each respondent with inter-wave activity history as well

as the lifetime employment and occupation histories collected at Waves 2 and 3, respectively.

4. Employees have on average five years of tenure in their current job. This value is not out of line with estimates of current job tenure in Britain found in other studies using different data sources (Mumford & Smith, 2004).

5. Age, age squared, work experience, experience squared, current job tenure, tenure squared, marital status, gender, having a dependent child, part-time employment, permanent contract, trade union membership, formal education received, having a vocational qualification, job leaver, promoted with current employer.

6. Industrial sector, workplace size, private ownership, non-profit and region.

7. The overall test of the explanatory power of the regressors is significant at a 99% confidence level for all the regressions and whilst the goodness of fit measures are not high, they are comparable with those found in other studies of training (see Leuven, 2002). Full results are available from the authors on request.

8. The set of interaction terms considered in the model and found to be statistically insignificant are: training × years of school; training × female; training × tenure; training × tenure^2; training × part-time; training × log hours; training × promoted and training × several occupation measures reveal. The inclusion of a quadratic term for training is statistically insignificant (at a level of 15%) and equal to zero.

9. The estimates for training are robust to the inclusion of the promotion measure in the set of explanatory variables. Nevertheless, promotion has a significant and a positive relationship with wages. Employees can expect their wage to rise by 4% when they are promoted.

10. The returns to training (incidence or intensity) are not found to be significantly different for public and private sector employees in any of the models considered in this chapter. These results are available from the authors upon request.

11. For example, Booth and Bryan (2007) use a subset of recent training occasions, the three longest in each year, which they divide into non-mutually exclusive current job training categories.

12. This finding may be inconsistent with the predictions of recent non-competitive models but still consistent with classical human capital theory in the presence of long-term labour contracts (Lazear & Oyer, 2004).

13. Given the imprecise nature of the blue-collar training estimates, a full set of white-collar and training interactive variables were introduced in the pooled sample (of all employees) to establish if the relationship between training and earnings is significantly different for blue- and white-collar employees. For all the training categories considered, the white-collar returns are found to be significantly different from those of the blue-collar employees at a minimum 80% confidence level (full results available from the authors upon request).

14. When the group of blue-collar workers is taken as the standard competitive ($\Omega = 0$) the portion of the measured wage gap due to coefficients differentials is smaller and the portion due to endowments differentials larger compared to using the white-collar wage structure ($\Omega = 1$). Even in this case, however, most of the wage differential is explained by measured productivity differentials across white- and blue-collar workers.

15. A limitation with the original Oaxaca (1973) approach is that the wage gap is measured at the mean, thereby ignoring potential differences in the form of the entire wage distribution. The use of quantile regressions allows for the decomposition of the wage gap at different points of the wage distribution. We explored the relationship between wages and training (for all three of our training measures) using quantile regression techniques and did not find significant differences across the wage distribution. In our particular example, where we are interested in a comparison of high- and low-skill workers (rather than higher and lower waged workers, see Chzhen & Mumford, 2009) we believe that the Oaxaca decomposition continues to be a valid and a pertinent approach.

ACKNOWLEDGMENT

Almeida-Santos is grateful for funding from the Fundacao para a Ciencia e Tecnologia-Ministerio da Ciencia e Tecnologia (Portugal).

REFERENCES

Acemoglu, D., & Pischke, J. S. (1999). Beyond Becker: Training in imperfect labor markets. *The Economic Journal*, *109*, F112–F142.

Almeida-Santos, F., & Mumford, K. (2005). Employee training and wage compression in Britain. *The Manchester School*, *73*(3), 321–342.

Ariga, K., & Brunello, G. (2006). Are the more educated receiving more training? Evidence from Thailand. *Industrial and Labor Relations Review*, *59*(4), 613–629.

Arulampalam, W., Booth, A., & Bryan, M. (2004). *Are there asymmetries in the effects of training on the conditional male wage distribution?* IZA Discussion Paper no. 984.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, *70*, 9–49.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education* (3rd ed.). Chicago, IL: The University of Chicago Press.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy*, *75*, 352–365.

Bishop, J. H. (1997). What we know about employer-provided training? A review of the literature. *Research in Labor Economics*, *16*, 19–87.

Booth, A. L., & Bryan, M. L. (2007). Who pays for general training in private sector Britain? *Research in Labor Economics*, *26*, 83–121.

Chiswick, B. (2003). *Review of Economics of the Household*, *1*(4), 343–362.

Chiswick, B., & Mincer, J. (1972). Time series changes in income inequality in the United States since 1939, with projections to 1985. *Journal of Political Economy (Supplement)*, *80*(2), S34–S66.

Chzhen, Y., & Mumford, K. (2009). *Decomposing gender gaps across earnings distributions in Britain*. Mimeo, University of York, UK.

Clifton, J. (1997). *Constraining influences on the decision to participate in training: the importance of the non-work environment*. Working Paper no. 97-25, Cornell-Center for Advanced Human Resource Studies, Ithaca, NY.

Cotton, J. (1988). On the decomposition of wage differentials. *Review of Economics and Statistics*, *70*, 236–243.

Department of Trade and Industry. (2005). Fairness at work. Chapter two. Business at work, retrieved from http://www.dti.gov.uk/er/fairness/part2.htm on 21/12/2005.

Frazis, H., & Loewenstein, M. A. (2005). Reexamining the returns to training: Functional form, magnitude and interpretation. *Journal of Human Resources*, *40*(2), 435–452.

Gershuny, J. (2005). Busyness as the badge of honor for the new super ordinate working class. *Social Research*, *72*(2), 287–314.

Halpin, B. (2006). *BHPS work-life history files, version 2*. Mimeo, ISER, University of Essex, Colchester. Available online at UKDA (documentation for study 3954).

Keep, E., Mayhew, K., & Corney, M. (2002). *Review of the evidence on the rate of return to employers of investment in training and employer training measures*. SKOPE Research Paper no. 34 (Summer), University of Warwick, UK.

Lazear, E., & Oyer, P. (2004). Internal and external labor markets: A personnel economics approach. *Labour Economics*, *11*(5), 527–554.

Leuven, E. (2002). *The economics of training: A survey of the literature*. Mimeo, retrieved from http://www.fee.uva.nl/scholar/mdw/leuven/reviewart.pdf

Leuven, E. (2004). A review of the wage returns to private sector training. *EC-OECD Seminar on Human Capital and Labour Market Performance*, Brussels.

Loewenstein, M. A., & Spletzer, J. R. (1998). Dividing the costs and returns to general training. *Journal of Labor Economics*, *16*(1), 142–171.

Lynch, L. M. (1992a). *Differential effects of post-school training on early career mobility*. NBER Working Paper Series no. 4034.

Lynch, L. M. (1992b). Private sector training and the earning of young workers. *American Economic Review*, *82*(1), 299–312.

Melero, E. (2004). *Evidence on training and career paths: Human capital, information and incentives*. IZA Discussion Paper no. 1377.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, *66*(4), 281–302.

Mincer, J. (1962). On-the-job training: Costs, returns and some implications. *Journal of Political Economy*, *70*(5, Part 2), S50–S79.

Mincer, J. (1970). The distribution of labor incomes: A survey with special reference to human capital approach. *The Journal of Economic Literature*, *VII*(March), 1–26.

Mincer, J. (1974). *Schooling, experience and earnings*. New York: Columbia University Press.

Mumford, K., & Smith, P. N. (2004). Job tenure in Britain: Employee characteristics versus workplace effects. *Economica*, *71*, 275–298.

Neumark, D. (1988). Employer's discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources*, *23*(3), 279–295.

Oaxaca, R. L. (1973). Male–female wage differentials in urban labor markets. *International Economic Review*, *14*(3), 693–709.

Oaxaca, R. L., & Ransom, M. R. (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics*, *61*, 5–24.

Ok, W., & Tergeist, P. (2002). Supporting economic growth through continuous education and training – Some preliminary results. Papers presented at the meeting of National Economic Research Organisations, OECD headquarters, Paris.

Pischke, J. S. (2001). Continuous training in Germany. *Journal of Population Economics*, *14*, 523–548.

Polachek, S. (1975). Differences in expected post-school investment as a determinant of market wage differentials. *International Economic Review*, *16*(2), 451–470.

Polachek, S. (2003). Mincer's overtaking point and the lifecycle earnings distribution. *Review of Economics of the Household*, *1*(4), 273–304.

Polachek, S. (2008). Earnings over the lifecycle: The Mincer earnings function and its applications. *Foundations and Trends in Microeconomics*, *4*(3), 165–272.

Regan, T. L., & Oaxaca, R. L. (2009). Work experience as a source of specification error in earnings models: Implications for gender wage decompositions. *Journal of Population Economics*, *22*(2), 463–499.

Reimers, C. (1983). Labor market discrimination against Hispanic and black men. *Review of Economics and Statistics*, *65*, 570–579.

Schøne, P. (2004). Why is the return to training so high? *Labour*, *18*(3), 363–378.

INCOME INEQUALITY, INCOME MOBILITY, AND SOCIAL WELFARE FOR URBAN AND RURAL HOUSEHOLDS OF CHINA AND THE UNITED STATES

Niny Khor and John Pencavel

ABSTRACT

In the United States, there is little difference in annual income inequality and income mobility between the rural and urban sectors of the economy. This forms a sharp contrast with China where income inequality is greater and income mobility lower among rural households than among urban households. When incomes are averaged over three years and when adjustments are made for the size and composition of households, income inequality among all *households differs little between China and the United States in the 1990s. Moreover when pooling rural households and urban households and when measuring annual income inequality and income mobility of the pooled households, the mobility of incomes of households in the United States differs little from that in China. Social welfare functions are posited that allow for a trade-off between increases in income and increases in income inequality. These suggest strong increases in well-being for urban households in China. The corresponding*

Jobs, Training and Worker Well-Being
Research in Labor Economics, Volume 30, 61–106
Copyright © 2010 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 0147-9121/doi:10.1108/S0147-9121(2010)0000030006

changes in rural China and in the United States are smaller. Four sets of data on households are drawn on to document these findings.

1. INTRODUCTION

The distinction between the urban and rural sectors of an economy has been a key feature of many models of economic development. Reflecting productivity differences of the activities in the two sectors, the central tendency of rural incomes tends to be lower than that of urban incomes. These income differences form the basis of models of rural–urban labor migration.

This focus on the central tendency of incomes neglects the fact that the income distributions in the two sectors often overlap considerably. One expression of these urban–rural differences in annual income is provided by Fig. 1, which graphs the frequency distribution of household income in China in 1995 among rural and urban households separately. Manifestly the distribution is displaced to the right among urban households compared with that for rural households. In addition, the income distribution appears

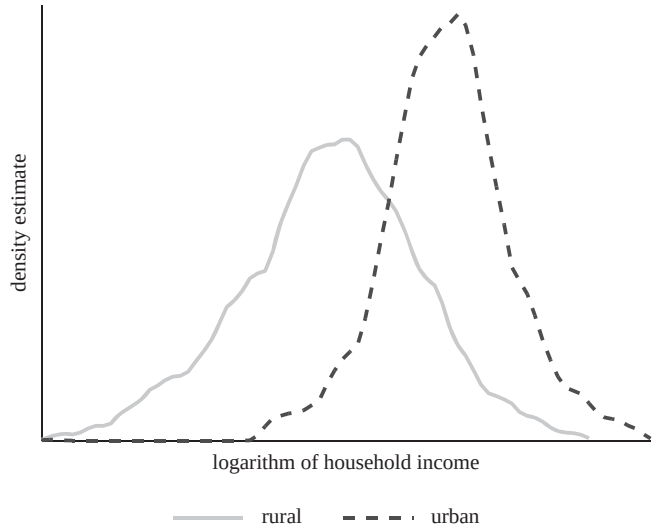

Fig. 1. Kernel Density Estimates of the Frequency Distribution of the Logarithm of Household Income 1995: Rural and Urban Households, China.

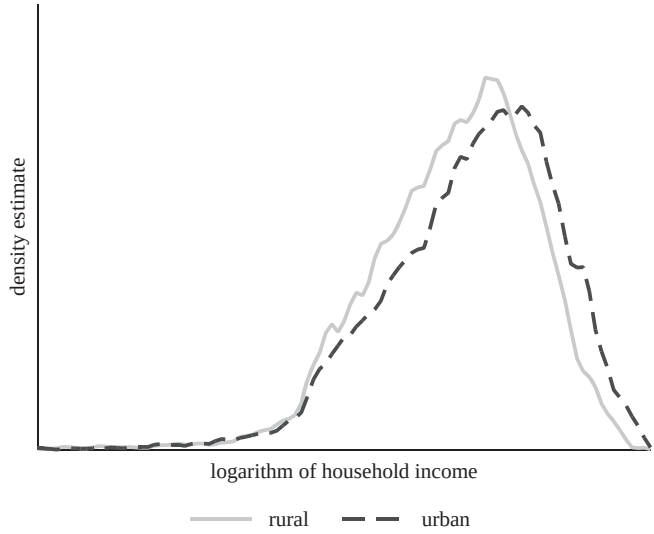

Fig. 2. Kernel Density Estimates of the Frequency Distribution of the Logarithm of Household Income 1995: Rural and Urban Households, United States.

narrower for urban households. The pattern is qualitatively the same in the United States as shown in Fig. 2: the rural household income distribution is to the left of that in urban areas. However, the degree of displacement of the rural relative to the urban income distribution in the United States is considerably less than that in China.[1] In addition, unlike China, it is not evident that the dispersion of rural incomes in the United States is different from the dispersion of urban incomes. So, while an urban–rural gap in household income is present in both China and the United States, its magnitude and other features of the income distribution appear quite different.[2] The degree to which the income distributions overlap is apparent in both countries.

This chapter is concerned with describing and analyzing the distribution of incomes in the rural and urban sectors of two economies, an emerging economy, China, and a developed economy, the United States of America. The gap in China between rural and urban incomes has been the subject of much research and related policy debate. It is useful, if not essential, to place the facts on urban–rural income differences in China in a comparative context and a contrast with a modern mature economy such as the United States provides an appealing perspective. This is particularly the case in view of the abiding issue of the degree to which different degrees of income

inequality are linked to alternative systems of economic organization. Whether Capitalism generates more enduring inequalities has been a major question of social analysis for at least well over 150 years. The facts regarding income inequality in modern China and the United States would appear to bring some empirical light to this issue especially as China has been moving away from a state-directed planned economy and toward a more decentralized market economy.

In making these comparisons, it is important to recognize that the conventional use of annual incomes may provide a misleading indicator of inequality insofar as one society is characterized by more year-to-year change in economic status than the other.[3] Hence an important aspect of our analysis is to use the observations on income of the same households over several years to determine the degree to which differences in income inequality in a given year are ameliorated through income mobility over time. How is the rural–urban difference in *annual* income inequality affected when using income information over several years for the same households? Are there important differences between rural households and urban households in the degree of income inequality and income mobility?[4] Four sets of data (two for China and two for the United States) are drawn upon to address these questions.

In addition, this chapter asks how we evaluate a situation in which incomes are growing at different rates among rural and urban households and, simultaneously, income inequality is changing. Insofar as society is averse to income inequality, what is the trade-off between increases in income and increases in income inequality? Of course, the answer to this question will depend critically on attitudes toward inequality, but the economist can provide a representation in which these values are given some quantitative expression. This is our task.

2. DATA SOURCES AND PROCEDURES

2.1. China: Chinese Household Income Project (CHIP)

In earlier drafts of this chapter, information on household income in China was drawn exclusively from CHIP (Riskin, Zhao, & Li, 2000), which in 1996 surveyed about 8,000 rural households and almost 7,000 urban households.[5] The data are obtained from larger samples designed by National Bureau of Statistics of China (NBSC) though the questions on income differ from the NBS's surveys. Nonresponse is unusual although the urban sample excludes

those lacking a formal certificate of residence (*hukou*), an exclusion of growing importance as this population increases over time.[6]

The survey has a different design in rural from urban areas. Measures of income include not only cash payments but also income in kind, state-financed subsidies, and the consumption of agricultural products by households engaged in agricultural production. The income concept here is pretransfer/pretax household income (though some cash transfers are included). This is discussed in the appendix where it is compared with an alternative income concept that incorporates all transfers. Though particular results will depend on the concept of income employed, our investigation into the effects of changes in income definitions suggests that our principal findings are robust with respect to alternative definitions of household income. Households are asked to keep a record of their incomes and expenditures.

The 1996 survey is based on an earlier survey conducted in the spring of 1989 for the reference year 1988 (Griffin & Zhao, 1993; Khan & Riskin, 1998) and we compare our measures of income dispersion in 1995 with those in 1988.[7] The 1988 survey asks about income in a typical month and this is simply converted to annual income by multiplying by 12. In 1995, information on annual income was solicited. Details about the formation of our samples from these surveys are outlined in the appendix. All income information from CHIP is reported in 1995 yuan by applying the consumer price index as a deflator.[8] Throughout, to mitigate the impact of measurement errors that are most likely to be present in outlying values, we habitually trim the data by omitting the 0.5 percent of the lowest and the 0.5 percent of the largest values of income in any sample. Of course, this will reduce measures of income inequality that draw on information throughout the income distribution. When we assessed the impact of this trimming procedure, we found it had inconsequential effects on our important inferences about inequality.[9]

An important part of this chapter consists of the analysis of incomes of the same households over time. From CHIP, the source for this information consists of questions that, in the urban survey, asks respondents to provide their "total income" not only in 1995 but also for each year from 1990 to 1994. Respondents are instructed to examine their records before providing income information for earlier years. As already noted, the rural survey in China was designed a little differently and the retrospective information on income asks for income not in every year from 1990 to 1995 but for the years 1991, 1993, and 1995. Hence, in our analysis of this retrospective income information, for urban and rural households together, we are obliged to use

data for the three years 1991, 1993, and 1995. To remove obvious errors in this retrospective information, for each household, we examined the values of the observations over time and attempted to "clean" the data by applying procedures sketched in the appendix.

Though there are transfers such as food stamps and public housing benefits in the United States, such noncash income represented a larger fraction of income in China in the early 1990s than in the United States. These subsidies were tilted in favor of urban households in China especially in the case of the housing subsidy. The housing subsidy amounted to an average of one-third of total cash income received by urban households in 1988 while food subsidies averaged 10.7 percent of household cash income. In the early 1990s, these subsidies were drastically reduced so that, by 1995, their inclusion in or exclusion from the income distribution mattered far less. To assess the impact of adding subsidies to our measure of pretax/pretransfer income, because information on subsidies is not available for 1990, for urban households who enjoyed more of these subsidies, we imputed the amount of subsidies received by households in 1990 using data for the first round of CHIP in 1988.[10] The inclusion of subsidies slightly lowers income inequality measures as well as income mobility figures although the magnitude of the fall in income mobility is slight.

2.2. China: China Health and Nutrition Survey (CHNS)

Even though most information about incomes in *all* surveys is retrospective and even though we have gone to considerable lengths to remove detectable errors, there remains the issue of the extent to which reporting errors in CHIP drive our results. Therefore, after the results using CHIP were derived, we turned to a different source of information, the CHNS, to assess whether our principal findings for China are replicated in these panel data.[11] The 1991, 1993, and 1997 waves of the CHNS were used to trace the evolution of household income in rural and urban areas of nine Chinese provinces. As with the CHIP data, the CHNS income data are trimmed by deleting the lowest and highest 0.5 percent of values in any year. CHNS incomes are deflated using a province-specific price index so that all incomes are expressed in 1988 yuan in Liaoning. The principal purpose of using the CHNS data is to provide an independent source of information about the same households over time.

2.3. United States of America

For the United States, we draw on information on household income recorded in two sources: the Annual Demographic Files of the Current Population Survey (CPS) for March 1989 and March 1996; and the Panel Study of Income Dynamics (PSID). The PSID's methods for the coding of wage income were revised in 1993 which frustrates following the incomes of the same households before and after 1993 so we choose a period during which the income definitions were unaltered, namely, the surveys from 1994 to 1999 that relate to the years 1993–1998. Analogous to the years 1991, 1993, and 1995 for CHIP and 1991, 1993, and 1997 for the CHNS in China, we study the years 1994, 1996, and 1998 in the United States. In the PSID, we allocate households to rural and urban sectors based on their residence in 1998.[12] Of the urban households in 1998, 97 percent were also in urban areas five years earlier and, of the rural households in 1998, 81 percent were in rural areas five years earlier. Therefore, the US data embody some migration between urban and rural areas although such change characterizes a relatively small fraction of these households and the vast majority of these US households maintain their rural or urban identification over these five years.

The incomes of the US households are expressed in 1996 dollars by applying the personal consumption expenditures price deflator. Insofar as possible, we use the same definitions for the US data as those that relate to the Chinese surveys. To our sample of 3,673 US households, we apply PSID sample weights to derive a sample that reflects the US population.[13]

Such sample weights do not exist in the Chinese surveys. In CHIP, the risk that the sample does not fully reflect the Chinese population is made more serious because not all households with income information for 1995 are represented with their income data for 1991 and 1993. In part this was because some rural households were not asked for their income in years prior to 1995, but in other instances, presumably, the problem is one of nonresponse. Among 7,997 rural households with income data for 1995, there is usable income information in 1991 and 1993 also for 72 percent of them.[14] Among 6,932 urban households with income information in 1995, income information for 1991 and 1993 is available for 92 percent.

This problem of missing data poses the same sort of concern as the problem of attrition in panel data: when observations are missing nonrandomly, the sample of households with income information in all years is not representative of all households. To help evaluate this, we may determine if the households with missing income data from CHIP for 1991

and 1993 in China are systematically different from all the households who provided income information in 1995. To this effect, define a variable, Q, that takes the value of unity for a household in China with income information for all years (1991, 1993, and 1995) and of zero otherwise. Express Q as a function of a number of variables from the 1996 survey including the household's income in 1995 to determine whether those households without income information in 1991 and 1993 are drawn randomly from all parts of the 1995 income distribution. The relationship is computed by conventional logistic maximum likelihood methods and Table 1 reports the estimated effects of differences in the right-hand side variables on the probability of complete income information.[15]

Among both urban and rural households in China, the coefficient estimates attached to the income decile dummy variables suggest that the largest differences are associated with the richest households in 1995: for

Table 1. Marginal Effects from Logit Estimates of the Probability of Providing Income Information in All Three Years in CHIP: China.

	Rural Households	Urban Households
Age	0.0171 (0.0030)	0.0037 (0.0020)
(age squared)/100	−0.0015 (0.0003)	−0.0043 (0.0020)
Woman	−0.038 (0.026)	−0.008 (0.007)
Communist Party	0.019 (0.014)	0.004 (0.007)
Ethnic minority	−0.143 (0.023)	0.016 (0.014)
Schooling1	0.129 (0.097)	−0.006 (0.015)
Schooling2	0.058 (0.073)	0.017 (0.012)
Schooling3	−0.033 (0.043)	−0.005 (0.013)
Schooling4	0.015 (0.017)	−0.002 (0.012)
Schooling5	−0.003 (0.012)	−0.031 (0.017)
Schooling6	Reference	Reference
<10 percentile	Reference	Reference
10–20 percentile	0.007 (0.023)	0.021 (0.012)
20–30 percentile	0.007 (0.023)	0.019 (0.012)
30–40 percentile	−0.003 (0.023)	0.027 (0.012)
40–50 percentile	0.008 (0.023)	−0.004 (0.015)
50–60 percentile	0.021 (0.023)	0.009 (0.014)
60–70 percentile	−0.035 (0.024)	0.013 (0.013)
70–80 percentile	−0.043 (0.025)	−0.006 (0.011)
80–90 percentile	−0.116 (0.026)	−0.036 (0.018)
>90 percentile	−0.190 (0.027)	−0.114 (0.026)
No. of adults	−0.001 (0.005)	0.009 (0.004)
No. of children	0.005 (0.005)	0.016 (0.007)

households in the top income decile in 1995, the probability of providing complete income information is 19 percent in rural areas and 11 percent in urban areas below the probability in the lowest income decile. In urban areas, there is also the suggestion that complete income information is almost 3 percent lower in the lowest income decile in 1995 than in fourth income decile. Therefore, the sample from CHIP with complete income information does not appear to be entirely representative of all households and well-off households, in particular, are less likely to be included in the income data for all years. Consequently, indicators of income inequality for the year 1995 assume lower values for the sample of households with complete income information than for the entire sample. Because of problems of sample attrition and nonresponse, it is common for studies of long-run income inequality and income mobility to be conducted on samples of individuals or households that are not fully representative of the larger population. Although this is a frequent feature of research studies on these topics, it does not mean we may dismiss the seriousness of the potential problem that our inferences about China will be drawn from a sample not entirely representative of the population.

In addition to the problem of nonresponse, there is the problem of response error. The consequences of such measurement error on our measures of income mobility are difficult to assess without knowing the properties of the errors. Some results in the literature regarding measurement error in income are based on the presumption that measurement errors take the classical form, but measurement error in income is unlikely to be classical (Bound, Brown, & Mathiowitz, 2001; Hyslop & Imbens, 2001; Gottschalk & Huynh, 2006). Perhaps the most probable form of response error is that, independent of their true incomes, individuals report the same income (or the same fraction of income) in different years. If this occurs, this will suggest less change in the income distribution than is really the case and our measures of income mobility will provide a lower bound on true income mobility.

2.4. Household Size and Composition

In China, urban and rural households tend to be of different size and composition and these differences are not independent of household income. This is suggested by the data in Table 2 reporting the average number of children N^C, the average number of adults N^A, and the average number of people N^{A+C} for each income decile for rural and urban households in

Table 2. Household Size and Composition by Household Income Decile: China (CHIP, 1995 and CHNS, 1997) and the United States (1998), Rural and Urban Households.

	Income Deciles										
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	Mean
China, CHIP 1995, rural households											
N^C	1.08	1.11	1.25	1.32	1.38	1.41	1.35	1.34	1.39	1.26	1.29
N^A	2.87	2.76	2.72	2.87	2.96	3.08	3.17	3.20	3.27	3.55	3.05
N^{A+C}	3.95	3.87	3.98	4.19	4.34	4.50	4.53	4.54	4.66	4.82	4.34
China, CHNS 1997, rural households											
N^C	1.30	1.12	1.22	1.20	1.12	1.09	1.11	0.88	0.90	0.71	1.06
N^A	2.84	2.98	2.91	2.74	2.74	2.97	2.89	2.68	2.89	2.81	2.85
N^{A+C}	4.21	4.13	4.23	4.03	3.96	4.12	4.07	3.63	3.84	3.56	3.98
China, CHIP 1995, urban households											
N^C	0.72	0.75	0.79	0.75	0.73	0.72	0.69	0.60	0.56	0.56	0.69
N^A	2.10	2.25	2.25	2.32	2.38	2.40	2.47	2.63	2.75	2.92	2.45
N^{A+C}	2.82	3.00	3.04	3.07	3.11	3.12	3.15	3.23	3.31	3.48	3.13
China, CHNS 1997, urban households											
N^C	0.96	0.69	0.77	0.66	0.77	0.69	0.63	0.59	0.52	0.76	0.70
N^A	2.51	2.59	2.70	3.00	2.49	2.92	2.54	2.52	2.48	2.64	2.64
N^{A+C}	3.49	3.35	3.54	3.68	3.30	3.70	3.25	3.18	3.10	3.43	3.41
US 1998, PSID, rural households											
N^C	0.18	0.40	0.32	0.40	0.47	0.55	1.02	0.86	0.75	0.73	0.57
N^A	1.28	1.50	1.60	1.68	1.82	1.88	2.03	2.15	2.27	2.16	1.84
N^{A+C}	1.46	1.90	1.91	2.08	2.29	2.43	3.05	3.01	3.02	2.89	2.41
US 1998, PSID, urban households											
N^C	0.43	0.37	0.40	0.58	0.59	0.68	0.79	0.68	0.74	0.87	0.61
N^A	1.26	1.41	1.55	1.67	1.77	1.87	1.97	2.05	2.21	2.24	1.80
N^{A+C}	1.69	1.78	1.95	2.25	2.36	2.54	2.76	2.74	2.95	3.11	2.41

China and the United States. In China, rural households tend to be larger than urban households with, on average, rural households having about one-and-a-half times the number of children as those in urban households. Household size tends to be larger in higher income households though the link between income and household composition is different between urban and rural areas: the ratio of adults to children tends to be larger in urban areas at higher income levels than in rural areas.

In the United States, rural households are not larger than urban households. As in China, richer households tend to have more members

than poorer households. A larger fraction of US households consist of a single adult and these households tend to have a lower income than households with two adults living together.

To determine whether our inferences are independent of alternative ways of comparing different types of households, we invoked different adjustments for household size and composition. In addition to using total household income, y_i, with no adjustments for household size and structure, we computed per capita household income, $y_i/(N_i^A + N_i^C)$, and per equivalent adult household income defined as $y_i/(N_i^A + \theta.N_i^C)^v$ where θ is the weight attached to children and v the scale economies parameter. The implications of alternative values of θ and v were examined and our general inferences did not change noticeably with respect to different values chosen.[16] In the results below, we report per equivalent adult household using a value of 0.75 for θ and a value of 0.85 for v. These values imply that, for example, in evaluating the value of a given yuan or dollar of income, a household consisting of five adults and no children is "equivalent" to a household with two adults and four children.

2.5. Measures of Inequality

To measure income dispersion, in addition to the Gini coefficient, the ratio of income at the 90th percentile to income at the 10th percentile, the coefficient of variation of incomes, and the standard deviation of the logarithm of incomes, we present a measure of inequality based on the social welfare function approach to inequality.[17] We draw upon this research explicitly in Section 5 below where we assess the change in well-being in a society when the general level of incomes rises at a time of simultaneously increasing income inequality. For the present, we note the following expression to measure income inequality where m denotes the mean of incomes and n the number of households:

$$N_\varepsilon = 1 - \left[n^{-1} \sum_{i=1}^{n} \left(\frac{y_i}{m} \right)^{1-\varepsilon} \right]^{1/(1-\varepsilon)} \tag{1}$$

The computation of this expression requires the specification of the parameter ε: when ε is zero, the index N registers indifference to inequality and N_ε is zero; as ε assumes larger values so the index is more sensitive to incomes at the lower tail of the income distribution and N_ε increases in value.[18] Common values assumed for ε are between 0.5 and 2.

3. INCOME INEQUALITY AND MOBILITY AMONG RURAL AND URBAN HOUSEHOLDS

3.1. Annual Income Inequality

The first questions addressed concern the degree of income inequality in urban and rural areas and the degree to which any difference in income inequality measured on the basis of annual incomes is offset by differences in income mobility over time. If household income mobility is different in rural from urban areas, then inequality measured with incomes over a longer period than one year may be quite different from inequality measured with annual incomes. To examine this issue, we use first the income information from the 1996 Chinese Household Income Project on the 5,797 rural households and 6,357 urban households in China with data in all years. A visual representation of the frequency distribution of rural and urban household incomes in 1995 is provided by the kernel densities in Fig. 1 from which it is evident that, in China, the central tendency of urban incomes is above that of rural incomes. The difference in the logarithm of incomes at the median or the mean implies rural household income is about 43 percent of urban income.[19]

It is also evident from Fig. 1 that the annual income distribution among rural households in China is wider than that among urban households. This visual impression is confirmed by the indicators of income inequality in Table 3 from CHIP. Thus, the Gini coefficient of 1995 household income is

Table 3. Annual Per Equivalent Adult Household Income Inequality: Rural and Urban China in 1995 and 1997 and Rural and Urban United States in 1998.

	Rural Households			Urban Households		
	China, CHIP 1995	China, CHNS 1997	US, PSID 1998	China, CHIP 1995	China, CHNS 1997	US, PSID 1998
Gini coefficient	0.350	0.374	0.388	0.254	0.355	0.393
90th/10th % ratio	5.200	7.024	6.702	3.163	5.553	7.018
Coefficient of variation	0.721	0.718	0.830	0.485	0.627	0.845
SD of log income	0.665	0.750	0.786	0.459	0.713	0.838
Atkinson's N: $\varepsilon = 0.5$	0.100	0.113	0.125	0.051	0.094	0.130
Atkinson's N: $\varepsilon = 1.0$	0.192	0.224	0.242	0.100	0.193	0.257
Atkinson's N: $\varepsilon = 2.0$	0.361	0.430	0.534	0.190	0.406	0.645

0.350 for rural Chinese households and 0.254 for urban Chinese households. Whereas incomes at the 90th percentile are about three times incomes at the 10th percentile among urban households, they are well over five times among rural households. In general, the indicators of income inequality in urban areas of China are between one half and three quarters their corresponding values in rural areas.[20]

The corresponding figures for urban and rural households in the United States do not suggest the same pattern: in the United States, annual income inequality among urban households exceeds that among rural households. For almost every inequality indicator in Table 3, annual household income inequality in the United States exceeds that in China.[21] The inequality gap between the United States and China is greater among urban households than among rural households.

3.2. Indicators of Income Mobility

3.2.1. Income Quintiles

Is there a difference in income mobility between rural households and urban households? A familiar method to address this question is to construct income transition matrices. An income transition matrix cross-classifies households into income quintiles from I (the bottom or poorest quintile) to V (the top or richest quintile) in two years. Each quintile contains the same number of households.[22] Each element of the income transition table consists of p_{jk}, the fraction of households in income quintile j in one year occupying income quintile k in a subsequent year.

For China, the two years are 1991 and 1995 using the CHIP data and the years 1993 and 1997 using the CHNS data. The transition matrix for rural households in China is presented in Table 4 and the matrix for urban households in Table 5. A χ^2 test of the null hypothesis that the transition matrices are symmetric cannot be rejected with a high level of confidence.[23] Consider the income information from the CHIP data. According to the top panel of Table 4, in rural areas of China, 61 percent of those who occupied the poorest fifth of households in 1991 were in the same quintile in 1995 whereas, according to the top panel of Table 5, in urban areas of China, 48 percent of the poorest households in 1991 were still in the lowest income category in 1995. In other words, this particular element of the tables suggests more income mobility in urban China than in rural areas.

Or consider mobility among the richest households. According to the CHIP data, among rural households in China, 60 percent of those who

Table 4. Per Equivalent Adult Household Income Transition Matrix: Rural Households.

China, CHIP						
		Year 1995				
		I	II	III	IV	V
Year 1991	I	0.613	0.213	0.114	0.035	0.024
	II	0.242	0.361	0.236	0.118	0.043
	III	0.090	0.267	0.311	0.235	0.097
	IV	0.037	0.136	0.251	0.338	0.237
	V	0.017	0.022	0.089	0.274	0.599
China, CHNS						
		Year 1997				
		I	II	III	IV	V
Year 1993	I	0.332	0.253	0.191	0.140	0.084
	II	0.261	0.299	0.210	0.156	0.073
	III	0.148	0.218	0.232	0.259	0.143
	IV	0.170	0.162	0.107	0.210	0.262
	V	0.089	0.067	0.170	0.235	0.438
US, PSID						
		Year 1998				
		I	II	III	IV	V
Year 1994	I	0.617	0.205	0.114	0.025	0.018
	II	0.219	0.468	0.163	0.098	0.054
	III	0.115	0.195	0.416	0.243	0.050
	IV	0.046	0.076	0.225	0.444	0.218
	V	0.004	0.056	0.081	0.190	0.660

occupied the richest income quintile in 1991 remained in that same quintile in 1995 whereas, among urban households, 54 percent of those in the top income quintile in 1991 were in the same quintile in 1995. Again, there is a suggestion of greater income mobility in urban than in rural areas. This is also implied by the CHNS data although the rural–urban difference is smaller in these data.

The transition matrices for the United States between 1994 and 1998 are presented in the bottom panels of Tables 4 and 5. Among rural households

Table 5. Per Equivalent Adult Household Income Transition Matrix: Urban Households.

China, CHIP						
		Year 1995				
		I	II	III	IV	V
Year 1991	I	0.478	0.234	0.157	0.101	0.029
	II	0.294	0.256	0.212	0.157	0.081
	III	0.153	0.249	0.263	0.202	0.133
	IV	0.067	0.206	0.229	0.277	0.221
	V	0.007	0.055	0.139	0.263	0.537
China, CHNS						
		Year 1997				
		I	II	III	IV	V
Year 1993	I	0.345	0.303	0.134	0.143	0.071
	II	0.261	0.239	0.232	0.134	0.136
	III	0.246	0.183	0.225	0.211	0.136
	IV	0.070	0.190	0.246	0.232	0.264
	V	0.077	0.085	0.162	0.275	0.393
US, PSID						
		Year 1998				
		I	II	III	IV	V
Year 1994	I	0.692	0.205	0.098	0.026	0.021
	II	0.173	0.463	0.231	0.108	0.038
	III	0.066	0.214	0.433	0.208	0.069
	IV	0.041	0.076	0.176	0.453	0.244
	V	0.029	0.041	0.062	0.206	0.629

in Table 4, 62 percent in the lowest income quintile in 1994 are still in the same quintile in 1998 and 66 percent in the highest income quintile occupy the same quintile in 1998. Among urban households in the United States in Table 5, the corresponding percentages are 69 and 63 percent respectively. These two numbers for urban households are little different from the respective numbers for rural households, which suggests income mobility in rural areas is similar to that among urban areas of the United States.

Table 6. Income Mobility-Income Quintiles for Rural and Urban China, 1991–1995 and 1993–1997 and Rural and Urban United States, 1994–1998.

	Rural Households			Urban Households		
	China, CHIP	China, CHNS	US, PSID.	China, CHIP	China, CHNS	US, PSID
Average quintile move	0.765	1.176	0.671	0.970	1.178	0.649
Immobility ratio	0.444	0.302	0.522	0.362	0.287	0.534
Adjusted immobility ratio	0.835	0.681	0.853	0.743	0.682	0.865

To facilitate comparisons of income mobility, consider three summary indicators of income mobility exhibited in the transition matrices: first, the average quintile move; second, the fraction who remain in the same quintile, also called the "immobility ratio"; and, third, an "adjusted immobility ratio," namely, the fraction who remain in the same quintile plus the fraction who move one quintile.[24] The computed values of these three summary indicators of income mobility between 1991 and 1995 in China, between 1993 and 1997 in China, and between 1994 and 1998 in the United States for rural and urban households are reported in Table 6. Within China, the CHIP data suggest that income mobility is higher among urban households than among rural households: the average quintile move is higher for urban households and the immobility ratio and the adjusted immobility ratio are lower for urban households compared with rural households. The CHNS data point to small differences in income mobility within China. In the United States, the average quintile move is higher among rural households and the immobility ratio and the adjusted immobility ratio higher for urban households, all of which suggest greater income mobility among rural than among urban households. So the urban–rural difference in the United States is different from that in China: based on these indicators from the income transition matrices, income mobility is greater among urban households than among rural households of China and income mobility is greater among rural households than among urban households in the United States. The gap between urban and rural households is smaller in the United States than that in China.

Finally, in every comparison between China and the United States in Table 6, that is, comparing urban China with urban US and comparing rural China with rural US, there is more income mobility in China than in the United States. This is true both for the CHIP data and for the CHNS

data. The China–US gap is especially marked among urban households. This is consistent with earlier research that focused on urban households alone (Khor & Pencavel, 2006).

3.2.2. Income Clusters

The indicators of income mobility discussed in the previous paragraphs are not invariant to the extent of income inequality in a society. In other words, a household experiencing a given increase in income is more likely to cross quintiles in an economy with a narrow income distribution than a household experiencing the same income increase in a society with a wide income distribution. Because the inequality in the annual income distribution in the United States is different from that in China and because the inequality of the annual distribution of income is different in rural areas from that in urban areas, consider constructing an income transition matrix defined not on the basis of income quintiles but on the basis of deviations from median income.

To be specific, specify five income clusters as follows: the lowest cluster consists of households with less than 0.65 of the median income; the second cluster consists of households with incomes between 0.65 and 0.95 of the median income; the third income cluster consists of households with incomes between 0.95 and 1.25 of the median income; the fourth cluster consists of households with incomes between 1.25 and 1.55 of the median income; and the fifth cluster consists of households with incomes above 1.55 of the median income. Obviously, if the median is the same in the two societies, the income cutoffs will be the same, but they will correspond to different fractions of households when income dispersion is different in the two societies. In a society with a wide income distribution, more households will be in the income cluster of less than 0.65 of the median compared with a society with a narrow income distribution. Now, however, households experiencing a given absolute increase in income in two societies will be equally likely to cross the thresholds between income clusters.

The consequence for our indicators of income mobility in China of measuring transitions across income clusters rather than transitions across income quintiles is shown in Table 7. The difference in mobility between rural and urban areas of China attenuates: as expected, in rural areas of China where the income distribution is wider, mobility appears to be greater when measured by movements across income clusters than measured by movements across income quintiles; and, in urban areas where the annual income distribution is narrower, mobility tends to be less when measured by transitions across income clusters than measured by transitions across

Table 7. Per Equivalent Adult Household Income Mobility: Income Clusters for Rural and Urban China and the United States.

	Rural Households			Urban Households		
	China, CHIP 1991–1995	China, CHNS 1993–1997	US, PSID 1994–1998	China, CHIP 1991–1995	China, CHNS 1993–1997	US, PSID 1994–1998
Average cluster move	0.839	1.337	0.671	0.913	1.274	0.649
Immobility ratio	0.464	0.339	0.522	0.367	0.311	0.534
Adjusted immobility ratio	0.801	0.622	0.853	0.777	0.633	0.865

income quintiles. According to the CHIP data, household income mobility in urban areas of China exceeds that in the rural areas of China.

In the United States, the differences in income mobility reported in Table 6 based on income quintiles tend to narrow or are even reversed in Table 7 when based on income clusters. A general conclusion from Tables 6 and 7 for the United States is that income mobility among urban households is not sharply different from income mobility among rural households. Using income quintiles or income clusters as a means to measure income mobility over five years, mobility among households in the United States is decidedly lower than mobility among Chinese households – at least when households are assigned to rural and urban sectors separately. This holds either for China's CHIP data or for China's CHNS data.

3.2.3. Factors Associated with Income Mobility

The indicators of income mobility in Table 6 describe the amount of income mobility across income quintiles over five years, but they are silent about those attributes of households that are associated with upward or downward mobility. Moreover, one might think of income mobility as a property that requires to be measured not simply between one pair of years but between many pairs of years. Put differently, because there are transitory factors that operate in any given year, the "permanent" probability of upward or downward income mobility is not fully observed using information on only one pair of years. Thus, define π_i as a latent index of permanent income mobility of household i and suppose π_i is a linear function of observed characteristics of the household X_i and unobserved factors, u_i:

$$\pi_i = \beta X_i + u_i \tag{2}$$

where u_i is assumed to be distributed normally with zero mean and unit variance. This standardized normal assumption will give rise to the estimation of an ordered probit model.

Although permanent income mobility π_i is unobserved, a household's position in the elements of the income transition matrices between any two years provides information on the permanent mobility of this household. Based on whether a household occupies an element on the diagonal of an income transition matrix or above the diagonal or below the diagonal, define a new variable z_i with the following features: $z_i = 1$ for households occupying a cell below the main diagonal (i.e., for households experiencing downward mobility), $z_i = 2$ for households occupying a cell on the main diagonal of the income transition matrix (households experiencing no mobility), and $z_i = 3$ for households in a cell above the main diagonal of the income transition matrix (households experiencing upward mobility).[25] The relation between the observed variable z_i and the latent variable π_i is given as follows:

$$\begin{aligned} z_i &= 1 \quad \text{if} \quad \pi_i \leq 0 \\ z_i &= 2 \quad \text{if} \quad 0 < \pi_i \leq \gamma_1 \\ z_i &= 3 \quad \text{if} \quad y_2 \leq \pi_i \end{aligned}$$

where γ_1 and γ_2 are censoring parameters to be estimated jointly with β. The X variables consist of household size and the following characteristics of the head of household: gender, years of age (entered as a quadratic form), years of schooling, an ethnic minority, and, for China, membership in the Communist Party.[26] The implications of the maximum likelihood estimation of the β parameters of Eq. (2) for the marginal effects are given in Tables 8 and 9 for China and in Table 10 for the United States.[27]

In general, for both China and the United States, the magnitude of the marginal effect of a given variable on the probability of upward mobility is close to the negative of the effect of the same variable on the probability of downward mobility. This is consistent with the symmetry of the income transition matrices, as reported earlier. In China, the marginal effects are not the same in the urban and rural sectors: female-headed households tend to be more upwardly mobile in urban areas than male-headed households whereas no meaningful gender differences in mobility in rural areas are evident; ethnic minorities tend to be more downwardly mobile in rural areas than nonminorities but such differences are not apparent in urban areas; while larger households tend to be more upwardly mobile in rural areas, there is no relation between household size and mobility in urban areas of

Table 8. Marginal Effects from Maximum Likelihood Estimation of the Probability of Upward and Downward Income Mobility: Urban and Rural China, CHIP from 1991 to 1995.

	Prob(Downward Mobility)		Prob(No Mobility)		Prob(Upward Mobility)	
	Rural	Urban	Rural	Urban	Rural	Urban
Woman = 1	0.010	−0.029	−0.001	0.001	−0.010	0.027
	(0.026)	(0.011)	(0.002)	(0.001)	(0.025)	(0.011)
Years of schooling	−0.0044	−0.0116	0.0001	0.0005	0.0043	0.0111
	(0.0019)	(0.0017)	(0.0001)	(0.0002)	(0.0018)	(0.0016)
Minority = 1	0.042	−0.041	−0.004	−0.001	−0.038	0.042
	(0.023)	(0.024)	(0.003)	(0.002)	(0.019)	(0.026)
Communist = 1	−0.011	−0.048	0.001	0.001	0.011	0.047
	(0.014)	(0.011)	(0.002)	(0.001)	(0.014)	(0.011)
Age/10	−0.035	−0.200	0.001	0.008	0.034	0.192
	(0.031)	(0.034)	(0.001)	(0.003)	(0.031)	(0.033)
$(\text{Age})^2/1{,}000$	0.045	0.199	−0.001	0.008	−0.044	−0.190
	(0.030)	(0.030)	(0.001)	(0.003)	(0.030)	(0.030)
Household size	−0.017	0.001	0.001	−0.001	0.017	−0.001
	(0.004)	(0.006)	(0.001)	(0.001)	(0.004)	(0.006)

Table 9. Marginal Effects from Maximum Likelihood Estimation of the Probability of Upward and Downward Income Mobility: Urban and Rural China, CHNS from 1993 to 1997.

	Prob(Downward Mobility)		Prob(No Mobility)		Prob(Upward Mobility)	
	Rural	Urban	Rural	Urban	Rural	Urban
Woman = 1	0.041	−0.072	−0.017	0.008	−0.025	0.064
	(0.036)	(0.042)	(0.034)	(0.041)	(0.035)	(0.043)
Years of schooling	−0.003	−0.003	0.004	−0.002	0.001	0.005
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.0005)
Minority = 1	0.053	0.001	0.020	0.019	−0.073	−0.020
	(0.033)	(0.060)	(0.032)	(0.058)	(0.031)	(0.059)
Communist = 1	0.054	−0.004	0.122	0.001	−0.068	0.002
	(0.093)	(0.063)	(0.101)	(0.060)	(0.092)	(0.063)
Age/10	0.042	0.030	−0.034	0.031	−0.007	−0.060
	(0.070)	(0.107)	(0.068)	(0.101)	(0.072)	(0.107)
$(\text{Age})^2/1{,}000$	−0.018	0.016	0.040	−0.042	−0.022	0.025
	(0.067)	(0.095)	(0.065)	(0.090)	(0.069)	(0.095)
Household size	−0.007	−0.024	0.001	−0.015	0.005	0.039
	(0.008)	(0.014)	(0.008)	(0.013)	(0.008)	(0.014)

Table 10. Marginal Effects from Maximum Likelihood Estimation of the Probability of Upward and Downward Income Mobility: Urban and Rural United States from 1994 to 1998.

	Prob(Downward Mobility)		Prob(No Mobility)		Prob(Upward Mobility)	
	Rural	Urban	Rural	Urban	Rural	Urban
Woman = 1	−0.020	−0.050	−0.002	−0.011	0.022	0.061
	(0.032)	(0.014)	(0.004)	(0.005)	(0.035)	(0.018)
Years of schooling	−0.0029	−0.0018	−0.0001	−0.0002	0.0031	0.0021
	(0.0043)	(0.0022)	(0.0003)	(0.0003)	(0.0045)	(0.0025)
Minority = 1	0.045	0.003	−0.001	0.001	−0.044	−0.004
	(0.034)	(0.014)	(0.004)	(0.002)	(0.031)	(0.016)
Years of age/10	0.056	0.136	0.003	0.017	−0.058	−0.153
	(0.045)	(0.025)	(0.004)	(0.006)	(0.047)	(0.028)
$(\text{Age})^2/1{,}000$	−0.038	−0.092	−0.002	−0.012	0.040	0.103
	(0.040)	(0.020)	(0.029)	(0.004)	(0.040)	(0.030)
Household size	−0.028	−0.019	−0.001	−0.002	0.029	0.021
	(0.010)	(0.046)	(0.002)	(0.001)	(0.010)	(0.005)

China; in CHIP, though the probability of upward income mobility follows an inverted U-shape with respect to age in both rural and urban areas, it reaches a peak at an age for those about 11 years younger in rural than in urban areas. More years of schooling are associated in China with a greater probability of upward income mobility.[28]

Whereas the marginal effects of variables on the probability of upward and downward mobility appear different in urban and rural areas of China, the corresponding marginal effects in the United States in rural areas are similar to those in urban areas. In the United States, the sign of the marginal effect of each variable on the probability of upward or of downward income mobility is the same in rural and in urban areas, something that is not a feature of the Chinese households. In the United States, female-headed households tend to be more upwardly mobile than male-headed households especially in urban areas. Upward mobility is more likely as household size increases. Minorities tend to have a lower probability of upward mobility than others especially in rural areas of the United States. Whereas in China the probability of upward mobility rises with age at a decreasing rate, in the United States the opposite pattern is evident: the probability of upward mobility falls with age at a decreasing rate. Though the rate of decline is not the same among rural and urban households, the age at which upward mobility reaches a minimum is virtually the same in urban and rural areas

and, at over 70 years of age, the great majority of heads of households are younger than the minimum and on the declining part of the mobility–age relationship.

These results indicate the differences in the mobility patterns of rural households and urban households in China. The sharp rural–urban differences in levels of income are exhibited also in rural–urban differences in the factors associated with income mobility. The empirical regularities associated with income mobility among urban households are not the same as the empirical regularities among rural households. This rural–urban difference in China is also not replicated in the United States where rural–urban differences in income mobility are of much smaller moment. All in all, there is much more meaning to the rural–urban distinction in China than to the rural–urban distinction in the United States.

3.2.4. Income Mobility among Pooled Households

These comparisons of annual income mobility have maintained the distinction between rural households and urban households. How does income mobility in China compare with that in the United States if urban and rural households are pooled? The transition matrices for all households are given in Table 11 for China and the United States with summary indicators of mobility in Table 12. These summary indicators suggest that the United States still appears less mobile than China when using the CHNS data but the US–China difference largely disappears in the CHIP data.[29]

4. A LONGER PERSPECTIVE ON INCOME INEQUALITY

4.1. Inequality among Rural Households and among Urban Households

What is the relationship between measures of inequality based on income averaged over three years and those based on income in a single year? At least for one measure of inequality, namely, the coefficient of variation of incomes, a precise expression may be derived. Suppose we have observations on incomes for years r, s, and t. Though it is not difficult to generalize the expression below, suppose the income distribution in each of these three years is stationary.[30] Then the coefficient of variation of income averaged

Table 11. Per Equivalent Adult Household Income Transition Matrix: All Households.

China, CHIP 1991–1995						
		Year 1995				
		I	II	III	IV	V
Year 1991	I	0.702	0.242	0.042	0.012	0.002
	II	0.252	0.445	0.208	0.074	0.021
	III	0.040	0.251	0.360	0.244	0.106
	IV	0.005	0.056	0.303	0.379	0.256
	V	0.001	0.006	0.088	0.291	0.614
China, CHNS 1993–1997						
		Year 1997				
		I	II	III	IV	V
Year 1993	I	0.365	0.256	0.197	0.102	0.080
	II	0.250	0.291	0.227	0.162	0.070
	III	0.174	0.195	0.252	0.229	0.150
	IV	0.133	0.180	0.156	0.264	0.267
	V	0.078	0.078	0.168	0.244	0.434
US, PSID 1994–1998						
		Year 1998				
		I	II	III	IV	V
Year 1994	I	0.667	0.225	0.098	0.025	0.024
	II	0.200	0.443	0.234	0.102	0.041
	III	0.071	0.208	0.452	0.208	0.063
	IV	0.037	0.073	0.165	0.454	0.247
	V	0.025	0.050	0.051	0.210	0.624

over the three years, C, may be written as

$$C = C_r\left(\frac{1}{3}\right)[3 + 2(\rho_{rs} + \rho_{st} + \rho_{rt})]^{1/2} \tag{3}$$

where C_r is the coefficient of variation in income in a single year r and ρ_{jk} the correlation coefficient between incomes in years j and k. Eq. (3) expresses the

Table 12. Summary Indicators of Per Equivalent Adult Household Income Mobility: Income Quintiles for all Chinese Households and all US Households.

	China		US
	CHIP, 1991–1995	CHNS, 1993–1997	PSID, 1994–1998
Average quintile move	0.600	1.176	0.654
Immobility ratio	0.500	0.296	0.528
Adjusted immobility ratio	0.909	0.677	0.868

Table 13. Correlation Coefficients of Per Adult Equivalent Household Income for the Same Households Across Different Years: Rural, Urban, and Pooled Households in China (CHIP, 1991–1995 and CHNS, 1991–1997) and in the United States (PSID, 1994–1998).

China, CHIP			China, CHNS			US, PSID		
	1993	1995		1993	1997		1996	1998
Rural								
1991	0.824	0.701	1991	0.453	0.308	1994	0.564	0.566
1993	1	0.765	1993	1	0.377	1996	1	0.590
Urban								
1991	0.877	0.643	1991	0.338	0.122	1994	0.877	0.643
1993	1	0.760	1993	1	0.310	1996	1	0.760
Rural and urban pooled								
1991	0.910	0.768	1991	0.421	0.258	1994	0.678	0.592
1993	1	0.846	1993	1	0.374	1996	1	0.668

inequality of income averaged over three years, C, as proportional to income inequality in a single year, C_r, where the factor of proportionality depends on the correlation coefficients in incomes, the values of ρ_{jk}. To help understand Eq. (3), consider limiting cases.

Suppose the correlation coefficients, ρ_{jk}, are all unity, complete income immobility. Then the factor of proportionality is unity and C equals C_r. As the correlation coefficients fall in value, so C falls relative to C_r. When all values of ρ_{jk} are zero, C is 58 percent of C_r and it requires negative values of ρ_{jk} to reduce C further as a fraction of C_r. In fact, the values of ρ_{jk} for China and the United States are given in Table 13. Whereas in the United States correlation coefficients are higher among urban than among rural

households, suggesting less income mobility among urban households, in China a consistent difference between rural and urban households is less apparent. However, with the CHIP data, for those correlation coefficients four years apart, 1991 and 1995, the correlation coefficients are higher among rural households than among urban households in China, consistent with the earlier result of greater income mobility between 1991 and 1995 in urban areas.

Using average values for ρ_{jk} for China and for the United States in Eq. (3) suggests that, in the CHIP data, inequality in the average of three year income will be about 93 percent of income inequality in a single year for China and about 88 percent of income inequality in a single year for the United States.[31] Indeed, according to Table 14, in China, inequality over three years of income is between 90 and 95 percent of inequality measured with incomes for 1995 alone and this figure is similar in rural and urban areas. In the United States in Table 14, inequality formed from incomes averaged over three years is between 85 and 90 percent of inequality based on 1995 incomes alone with small differences between urban and rural households.[32] The usefulness of Eq. (3) as a guide to thinking about the effect on measures of inequality of averaging over incomes in a number of years is evident.

4.1.1. Income Inequality among Pooled Households

These comparisons of rural and urban household incomes in China and the United States leave unanswered the question of the relative amount of income inequality in the two countries when rural households are combined with urban households. Which of these two societies manifests greater income inequality when all households are considered? Table 15 presents the values of various indicators of household inequality computed for the urban and rural households pooled. When using incomes in a single year, on most indicators, China reveals less income inequality than the United States.[33] This gap is attenuated when inequality is measured using incomes averaged over three years.

5. MEASURES OF CHANGES IN SOCIAL WELL-BEING

In this section, a social welfare function is exploited to trade-off changes in mean income and changes in income inequality in China and the United

Table 14. Per Equivalent Adult Household Income Inequality after Averaging Income over Years: China and the Unites States, Rural and Urban Households.

	China, CHIP			
	Rural Households		Urban Households	
	1995	1991, 1993, and 1995	1995	1991, 1993, and 1995
Gini coefficient	0.350	0.332	0.254	0.242
90th/10th % ratio	5.200	4.666	3.163	3.011
Coefficient of variation	0.721	0.669	0.485	0.461
SD of log income	0.665	0.625	0.459	0.435
Atkinson's N: $\varepsilon = 0.5$	0.100	0.090	0.051	0.046
Atkinson's N: $\varepsilon = 1.0$	0.192	0.173	0.100	0.091
Atkinson's N: $\varepsilon = 2.0$	0.361	0.327	0.190	0.173
	China, CHNS			
	Rural Households		Urban Households	
	1997	1991, 1993, and 1997	1997	1991, 1993, and 1997
Gini coefficient	0.374	0.301	0.355	0.251
90th/10th % ratio	7.024	4.151	5.553	3.673
Coefficient of variation	0.718	0.570	0.627	0.453
SD of log income	0.750	0.552	0.713	0.484
Atkinson's N: $\varepsilon = 0.5$	0.113	0.071	0.094	0.051
Atkinson's N: $\varepsilon = 1.0$	0.224	0.139	0.193	0.103
Atkinson's N: $\varepsilon = 2.0$	0.430	0.262	0.406	0.208
	US, PSID			
	Rural Households		Urban Households	
	1998	1994, 1996, and 1998	1998	1994, 1996, and 1998
Gini coefficient	0.388	0.359	0.393	0.362
90th/10th % ratio	6.702	5.529	7.018	6.218
Coefficient of variation	0.830	0.732	0.845	0.708
SD of log income	0.786	0.689	0.838	0.727
Atkinson's N: $\varepsilon = 0.5$	0.125	0.104	0.130	0.108
Atkinson's N: $\varepsilon = 1.0$	0.242	0.202	0.257	0.214
Atkinson's N: $\varepsilon = 2.0$	0.534	0.388	0.645	0.426

Table 15. Per Equivalent Adult Household Income Inequality: Urban and Rural Households Pooled China and the United States, Single Year and Average of Three Years.

	Single Year			Average of Three Years		
	China, CHIP, 1995	China, CHNS, 1997	US, PSID, 1998	China, CHIP, 1991, 1993, 1995	China, CHNS, 1991, 1993, 1997	US, PSID, 1994, 1996, 1998
Gini coefficient	0.387	0.368	0.397	0.375	0.313	0.367
90th/10th % ratio	7.546	6.875	7.268	6.955	4.630	6.210
Coefficient of variation	0.731	0.700	0.855	0.704	0.601	0.725
SD of log income	0.802	0.752	0.833	0.768	0.589	0.728
Atkinson's N: $\varepsilon = 0.5$	0.123	0.111	0.132	0.115	0.078	0.110
Atkinson's N: $\varepsilon = 1.0$	0.245	0.221	0.258	0.230	0.153	0.216
Atkinson's N: $\varepsilon = 2.0$	0.473	0.433	0.623	0.444	0.292	0.425

States. For this, panel data are not essential so we concentrate on the CHIP data for China and the CPS data for the United States.

5.1. China

According to the household income surveys for 1989 and 1996, average per equivalent adult household income increased in China between 1988 and 1995 by more than 5 percent per year. Table 16 indicates that the increase was considerably larger among urban than among rural households.[34] In part, this is because average household size fell in both rural and urban areas. This is shown in Table 17 where among rich households, in particular, the fall in average household size is quite remarkable: in the top deciles of annual income, on average among rural households, there was one fewer household member in 1995 than in 1988 and among urban households the drop is almost three quarters.

These increases in average household income in China were accompanied by a growth in annual income inequality as shown in Table 18. Between 1988 and 1995, the Gini coefficient increased among rural households from 0.295 to 0.350 and among urban households it increased from 0.207 to 0.254.[35] For all households, the Gini coefficient increased from 0.329 in 1988 to 0.387 in 1995. Other indicators of income inequality in China reveal similar increases in annual income inequality. In most instances, the increase

Table 16. Percent Annual Average Growth in Mean Real Per Equivalent Adult Household Income: China and the United States.

	Rural Households	Urban Households	Rural and Urban Households
China, CHIP, 1988–1995	2.377	5.433	5.154
US, CPS 1988–1995	0.773	0.614	0.803

Table 17. Change in Average Household Size by Household Income Decile, Urban and Rural China.

Income Decile										
1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	Mean
Rural, CHIP 1995–1988										
−0.06	−0.60	−0.70	−0.64	−0.63	−0.60	−0.74	−0.94	−0.96	−1.01	−0.69
Urban, CHIP 1995–1988										
−0.09	−0.24	−0.26	−0.36	−0.30	−0.42	−0.44	−0.55	−0.66	−0.73	−0.41

Table 18. Per Equivalent Adult Household Income Inequality in China: CHIP, 1988 and 1995.

	CHIP, 1988			CHIP, 1995		
	Pooled	Rural	Urban	Pooled	Rural	Urban
Gini coefficient	0.329	0.295	0.207	0.387	0.350	0.254
90th/10th % ratio	5.192	3.911	2.503	7.546	5.200	3.163
Coefficient of variation	0.615	0.582	0.403	0.731	0.721	0.485
SD of log income	0.646	0.540	0.376	0.802	0.665	0.459
Atkinson's N: $\varepsilon = 0.5$	0.088	0.070	0.035	0.123	0.100	0.051
Atkinson's N: $\varepsilon = 1.0$	0.175	0.135	0.069	0.245	0.192	0.100
Atkinson's N: $\varepsilon = 2.0$	0.260	0.255	0.134	0.473	0.361	0.190

in income inequality among rural households exceeds the increase among urban households.

With higher incomes and yet greater income inequality, as a society, was China better-off in 1995 than in 1988? The answer depends, in part, on society's aversion to income inequality: a society that is indifferent to

inequality will prefer a situation in which incomes are higher regardless of how these higher incomes are distributed. However, usually, people are not indifferent to increasing inequality and a given increase in income received by a poor household is regarded as constituting a larger increase in social well-being than the same increase in income enjoyed by a rich household. An indicator of well-being that embodies these relative preferences toward levels of income and their distribution is Atkinson's (1970) additive social welfare function

$$V = (n)^{-1} \sum_{i=1}^{n} (1-\varepsilon)^{-1} y_i^{1-\varepsilon} \tag{4}$$

where the parameter $\varepsilon \geq 0$ regulates the trade-off between levels of income, y, and the distribution of income and n denotes the number of households.[36] Using the concept of the equally distributed equivalent income, the measure of inequality implied by this function is given by Eq. (1) above and this allows us to rewrite V in the more transparent way:

$$V = (1-\varepsilon)^{-1}[(1-N_\varepsilon)m]^{1-\varepsilon} \tag{5}$$

which makes clear the substitution possibilities between income levels as summarized in mean income m and income inequality, N_ε.[37] From Eq. (5), solve for the mean income needed to attain a given level of welfare, V, when inequality (as measured by Eq. (1)) takes a particular value:

$$m = (1-\varepsilon)^{1/(1-\varepsilon)} V^{1/(1-\varepsilon)} (1-N_\varepsilon)^{-1}$$

Suppose we observe incomes in two periods, period s and period t, and set social welfare equal to the level enjoyed in period s. Then determine the value of m needed in period t to attain the level of well-being in period s given inequality in period t. Let m_t^V be this constant-welfare level of mean income in period t which, with this expression for social welfare, is given by the simple expression

$$m_t^V = \frac{(1-N_{\varepsilon s})}{(1-N_{\varepsilon t})} m_s$$

We may say social welfare has improved in period t over period s if mean income in period t, m_t, exceeds the level of income needed to maintain social welfare constant, m_t^V. In this way, an indicator of social welfare is derived, namely, m_t/m_t^V which allows for trade-offs between increases in the level of incomes and increases in income inequality. When m_t/m_t^V exceeds unity, social welfare in period t has improved relative to welfare in period s.

Table 19. Values of Indicators of the Change in Social Well-Being, m_t/m_t^V, in China.

	$\varepsilon = 0$	$\varepsilon = 0.5$	$\varepsilon = 1.0$	$\varepsilon = 1.5$	$\varepsilon = 2.0$
CHIP, 1988–1995					
Rural households	1.19	1.15	1.11	1.07	1.02
Urban households	1.47	1.44	1.42	1.40	1.37
Rural and urban households	1.44	1.39	1.32	1.24	1.15

Evidently, m_t/m_t^V depends on inequality in period s and inequality in period t and, because these measures of inequality depend on the inequality-aversion parameter ε, this indicator of the change in well-being m_t/m_t^V incorporates society's attitudes toward inequality. Thus, for a given value of ε, we have an expression that provides an index of the degree to which society was better-off (if at all) in China in 1995 with higher incomes that were distributed more unequally in 1995 than in 1988.

The values of m_t/m_t^V are listed in Table 19 for values of ε between 0 and 2 for rural households, urban households, and for the pooled (urban plus rural) households. When $\varepsilon = 0$, society is indifferent to inequality so, given the increase in income in China between 1988 and 1995 and because the increase in income inequality is disregarded, our welfare indicator should register the largest increase. Indeed, along any row of Table 19, the values of m_t/m_t^V are largest when $\varepsilon = 0$. Thus, when $\varepsilon = 0$, welfare is 19 percent higher in 1995 than in 1988 for rural households, 47 percent higher for urban households, and 44 percent higher for all households. However, as ε assumes larger values, so the increase in welfare is attenuated because the increase in income inequality between 1988 and 1995 assumes greater importance in our welfare indicator. In fact, when higher values of ε are posited, there is some doubt that welfare among rural households in China increased between 1988 and 1996. When $\varepsilon = 2$, welfare in urban areas in 1995 was 37 percent above that in 1988 and welfare for all households in 1995 was 15 percent above that in 1988, but welfare among rural households in 1995 was merely 2 percent above that in 1988. For a given ε, well-being rose more for urban households than rural households. This is because, as Table 16 indicates, real incomes increased more in urban areas than in rural areas and, as Table 18 shows, income inequality increased less in urban than in rural areas.

One may ask a slightly different question. Instead of positing a particular value for ε and then determining how welfare changed, one may ask what value of the inequality-aversion parameter is required for welfare in 1995 to be

Table 20. Values of ε Needed for Social Welfare in China Not to Have Increased between 1988 and 1995.

	Rural Households	Urban Households	Rural and Urban Households
CHIP between 1988 and 1995	2.20	33.51	2.92

the same as welfare in 1988. Expressed differently, given the actual changes in incomes, how much aversion to income inequality is required for social welfare not to have increased (i.e., for m_t/m_t^V to be unity).[38] The answers to this question are provided in Table 20 that highlights the different experiences of rural households and urban households. Given the income changes among urban households, a substantial inequality aversion of over 33 is needed to avoid the inference that social well-being between 1988 and 1995 did not increase. By contrast, among rural households, on the basis of total household income, an aversion to income inequality ε of a little over two is needed for social well-being in rural areas not to have increased. This underlines the sharp difference in the experience of urban and rural households.

5.2. The United States

The same analysis is now undertaken for the United States using information from the Annual Demographic files of the CPSs for March 1989 and March 1996 that relate to incomes in 1988 and 1995 respectively, the same period as that covered by the Chinese income data. As is evident from the lower panel of Table 16, the annual average growth in real per equivalent adult household income between 1988 and 1995 in the United States was considerably smaller than in China.[39] In addition to a meager growth in household income, Table 21 indicates that, for almost all measures of inequality, annual household income inequality grew in the United States between 1988 and 1995. Almost the only indicator of inequality not suggesting this is N_ε corresponding to $\varepsilon = 2$ for which, among rural households, income inequality narrows.[40]

Thus, with modest increases in average income and with most gauges suggesting increases in income inequality, the welfare indicator m_t/m_t^V for the United States is likely to register small or negative changes in well-being. Indeed, this is the suggestion of the values of m_t/m_t^V reported in Table 22: among both rural and urban households, when $\varepsilon = 0.5$ or 1.0, well-being in

Table 21. Per Adult Equivalent Household Income Inequality in the United States: 1988 and 1995.

	1988			1995		
	Pooled	Rural	Urban	Pooled	Rural	Urban
Gini coefficient	0.392	0.379	0.389	0.416	0.387	0.418
90th/10th % ratio	7.344	6.391	7.344	7.692	6.393	7.692
Coefficient of variation	0.781	0.754	0.769	0.904	0.795	0.908
SD of log income	0.858	0.828	0.858	0.903	0.827	0.919
Atkinson's N: $\varepsilon = 0.5$	0.128	0.119	0.126	0.144	0.124	0.146
Atkinson's N: $\varepsilon = 1.0$	0.258	0.241	0.257	0.284	0.247	0.289
Atkinson's N: $\varepsilon = 2.0$	0.834	0.902	0.737	0.939	0.838	0.950

Table 22. Values of Indicators of the Change in Social Well-Being, m_t/m_t^V, in the United States from 1988 to 1995.

	$\varepsilon = 0$	$\varepsilon = 0.5$	$\varepsilon = 1.0$	$\varepsilon = 1.5$	$\varepsilon = 2.0$
Per equivalent adult household income					
Rural households	1.05	1.04	1.04	1.08	1.73
Urban households	1.05	1.03	1.01	0.89	0.20
Rural and urban households	1.06	1.04	1.02	0.94	0.39

1995 is some 1–4 percent above that in 1988; for larger values of ε, except among rural households, well-being is lower in 1995 than in 1988 because the small increase in average income does not offset the increase in income inequality; the rural households are somewhat different because, as shown in Table 21, for $\varepsilon = 2$, $N\varepsilon$ suggests less, not more, income inequality and so, with higher average income and with less inequality, well-being in 1995 is considerably higher among rural households (when $\varepsilon = 2$) than in 1988.

Instead of hypothesizing a value of ε and assessing the change in welfare, determine the value of the inequality-aversion parameter ε such that welfare in 1995 is the same as welfare in 1988 in the United States. In other words, given the observed changes in incomes, how much aversion to income inequality is needed for social welfare not to have increased (i.e., for m_t/m_t^V to be unity). The answers to this question are provided in Table 23 that highlights the different experiences of rural households and urban households. Given the small income changes among urban households, only a

Table 23. Values of ε Needed for Social Welfare in the United States Not to Have Increased between 1988 and 1995.

	Rural Households	Urban Households	Rural and Urban Households
Per equivalent adult household income	4.19	0.99	1.30

modest aversion to inequality is needed to avoid the inference that social well-being between 1988 and 1995 did not increase.

6. CONCLUSIONS

The research in this chapter has reported on the distribution of incomes among households in rural and urban areas of a rapidly developing economy and of a mature economy, China and the United States, respectively, in the 1990s. The pattern of incomes in China and the United States is quite different. In China, annual income inequality is wider and annual income mobility is lower among rural households than among urban households. This is the case for two independent data sets. By contrast, there is little difference in the United States in annual income inequality and in income mobility between rural and urban households. Similarly, the variables associated with income mobility are not the same among rural households as among urban households in China whereas these variables have a similar association with mobility among rural as among urban households in the United States. Though these income differences between rural and urban households are greater in China than in the United States, during these years, when examining rural households and urban households separately, there tends to be less annual income inequality and greater income mobility in China than in the United States.

This conclusion – of greater annual income inequality and less income mobility in the United States than in China – holds when examining the annual incomes of rural households and of urban households separately. When incomes are averaged over three years (adjusting for the size and composition of households), income inequality among *all* households differs little between China and the United States in the 1990s. Moreover when pooling rural households and urban households and when measuring annual income inequality and income mobility of the pooled households, the

mobility of incomes of households in the United States differs little from that in China.

In both China and the United States, household incomes have tended to grow at a time when income inequality has widened. If societies are averse to income inequality, from a social welfare perspective, has the growth in incomes offset the increase in income inequality? We address this question for urban and rural households in China and the United States. The answer requires a judgment about society's values and, in particular, about the weight placed on income inequality in the expression of society's welfare. Using a metric for social welfare that compares actual incomes in 1995 with those incomes needed to maintain well-being the same as in 1988, we find unambiguous increases in social well-being for urban households in China where the strong rise in incomes clearly offsets the relatively small increases in income inequality. Among rural households in China, the modest increases in incomes were adequate to compensate for increases in income inequality only when society exhibits low levels of aversion to income inequality. For the United States, income growth was smaller than in China and, given the increases in income inequality suggested by most indicators of inequality, so the growth in social well-being in the United States was lower than in China. However, in the United States, changes in social well-being among urban households is similar to that among rural households (except for social welfare functions that are strongly averse to income inequality).

A persistent finding is that the urban–rural distinction embodies much greater meaning for households in China than for households in the United States. For the United States, the income inequality and mobility patterns among urban households are similar to those among rural households. This is not the case in China where the rural and urban sectors are much more distinct economies.

NOTES

1. Glaeser and Mare (2001) analyze the rural–urban difference in the central tendency of labor incomes in the United States.

2. The data for China are from the Chinese Household Income Project described below and those for the United States from the Annual Demographic File of the Current Population Survey for March 1996. The densities are estimated using the Epanechnikov kernel with a bandwidth of 0.05.

3. Friedman (1962, pp. 171–172) provides a robust statement of the argument that measures of *annual* income are especially ill-suited to assess inequality in Capitalist societies which are apt to more turbulent and mutable than Socialist societies.

4. There is little research addressing the issues in this paragraph for Chinese households using panel data. An earlier paper focused on income mobility among urban households only and, even for those urban households, did not take up the same set of questions (Khor & Pencavel, 2006).

5. The Chinese Household Income Project is a research effort jointly sponsored by the Institute of Economics, Chinese Academy of Social Sciences, the Asian Development Bank, and the Ford Foundation with additional support provided by the East Asian Institute, Columbia University. Khan and Riskin (2001) provide a careful analysis of some findings.

6. The 2002 survey includes information on those moving to urban areas without a *hukou*. See Deng and Gustafsson (2006) and Ximing, Sicular, Li, and Gustafsson (2008).

7. The third wave for 2002 includes a sample of migrants whose incomes tend to lie between those of urban and rural households (Khan, 2004).

8. In their comprehensive analysis of the 1988 and 1995 household income data, Khan and Riskin (2001) use the National Bureau of Statistics (NBS) consumer price index numbers to deflate rural incomes slightly differently from urban incomes. With 1988 = 100, the NBS's Rural CPI is 220.09 in 1995 and the Urban CPI is 227.90 in 1995. They express the suspicion that these price increases understate the amount of inflation over this time. We note the small difference implied in price inflation between rural and urban areas. The price deflator we use takes the value of 223.1 in 1995 with 1988 = 100. Benjamin, Brandt, and Giles (2005) compare movements in rural household inequality that deflate incomes with a spatially insensitive price index with those that use a price index that varies across provinces. In any year, the Gini coefficient is some 2 or 3 percent lower with the spatially sensitive price index but the movements over time in the Gini coefficient are very similar regardless of the price deflator. Démurger, Fournier, and Li (2005) also compare the effects on inequality indicators of using a provincial price deflator. For urban households in 1995, the Gini coefficient of per adult equivalent household disposable income without such deflation is 0.321 and is 0.298 when a province-sensitive price deflator is used. This difference is similar to that reported for rural households by Benjamin et al. (2005). The CHNS income data described below are deflated by a price index that varies by province and by rural–urban sector.

9. The measures for China of the central tendency of incomes, the dispersion of incomes, and income mobility that are presented from CHIP in this chapter for pooled urban and rural households together are unweighted by their selection probabilities because the surveys do not supply these. However, we created our own weights using population by provinces as weights and calculated descriptive statistics weighting by the reciprocal of these sampling probabilities. There was little difference between the weighted and the unweighted values and, to show this, we report some weighted values in footnotes below. Cowell, Litchfield, and Mercader-Prats (1999) provide an analysis and application of the practice of trimming the tails of income distribution data. The deletion of outliers is a standard (though by no means universal) procedure in labor economics. Card, Lemieux, and Riddell (2004) is a recent example that uses the Current Population Survey, as we do.

10. We calculated the level of the housing, food, and other subsidies from the 1988 CHIP for households with particular characteristics (such as attributes of the household head and geographic identifiers). We identified households with these

characteristics in the 1996 CHIP and, using the 1988 associations between these characteristics and subsidies, we imputed the subsidies for these households in 1990. Such imputed subsidy-augmented incomes in 1990 were compared with actual subsidy-augmented incomes in 1995 to determine the impact of including such subsidies on our inferences about income mobility. Housing subsidies were especially generous for urban households so we applied this imputation procedure for urban households only for whom the effect of inclusion or exclusion of subsidies in total income is probably more important.

11. We are by no means the first to make use of the household income data in the CHNS. For instance, in a paper that became known to us after the second draft of this chapter was completed, Fields and Zhang (2007) make use of both CHIP and CHNS data. Also Benjamin, Brandt, Giles, and Sangui (2008) use the CHNS as repeated cross-sections to describe changes in income inequality from 1991 to 2000. The CHNS is administered jointly by the Chinese Center for Disease Control and Prevention and the University of North Carolina Population Center. See http://www.cpc.unc.edu/projects/china.

12. Analogously, the location of the Chinese households in CHIP is determined by their residence in the final year, 1995. There was relatively little rural–urban movement of these households in China in the early 1990s except among those without hukou who are not covered by this household survey. The impact of hukou on mobility at this time is discussed in Deng and Gustafsson (2006). In the CHNS, none of our households reveals a change in urban–rural status until after 2000.

13. For the United States, the PSID provides information on the characteristics of the county in which the household resides and one of these characteristics is the area's population. For the results reported in this chapter, an urban household is one living in an area with a population greater than 20,000. This definition results in an urban population for the United States that constitutes 75 percent of the total and this compares with the Census Bureau's definition that allocated 79 percent of the US population in the 2000 Census to urban areas. See http://www.ers.usda.gov/Briefing/Rurality/WhatisRural/. We did investigate other allocations of areas between the rural and urban categories, but our inferences about income inequality were not affected to any material degree.

14. One problem with the rural CHIP file is a suspiciously large number of zero values for household income. Do these zeros really mean no household income or, more likely, was the information on income not recorded? In 1995, there are 11 households out of 7,997 with zero household income, there are 1,602 with zero income in 1993, and there are 2,060 with zero reported income in 1991. We have dropped all households reporting zero income from our analysis and this constitutes a major reason for why the 7,997 households in 1995 shrinks to 5,797 for our analysis sample (i.e., we work with 72 percent of the 1995 sample). As is well known, zero incomes may induce measurement difficulties for inequality indicators because some indicators are not well-defined or assume their limiting values in the presence of zeros (e.g., Atkinson's indicator with $\varepsilon = 1$ reaches its maximum value when incomes are zero). Issues concerning the interpretation and management of zero income values in surveys are addressed by Cowell et al. (1999).

15. In urban areas, there are 6,932 households with income data in 1995 and there is income information for 1991 and 1993 on 6,357 of them. In rural areas, of the

7,997 households with 1995 income data, there are 5,797 households with income information also in 1991 and 1993. In Table 1, estimated standard errors are in parentheses. For continuous variables, marginal effects are partial derivatives while, for discrete variables, the effects are of a change in the value of the dummy variable from zero to unity. These effects are evaluated at the mean values of the right-hand side variables. "Age" measures years of age of the head of household. "No. of adults" and "no. of children" are, respectively, the number of adults and number of children in the household (with someone 18 years or over constituting an adult). All the other variables are dichotomous variables. "Woman" takes the value of unity for a household headed by a woman, "married" takes the value of unity for a household head who is currently married. "Communist Party" takes the value of unity for a household head who is a member of the Communist Party and "ethnic minority" that takes the value of unity for a household head who reports being an ethnic minority. The schooling variables describe the years of schooling attained by the household head. "Schooling1" takes the value of unity for someone with a college education, "schooling2" takes the value of unity for someone with a professional school education, "schooling3" takes the value of unity for someone with a middle-level professional, technical or vocational school education, "schooling4" takes the value of unity for someone with an upper middle school education, and "schooling5" takes the value of unity for someone with a lower middle school education. "schooling6," the omitted category, refers to elementary or below elementary school. The variables taking the form "x–y percentile" are dichotomous variables that take the value of unity for a household with an income in 1995 in the percentile range between x and y. The lowest tenth percentile constitutes the reference category.

16. Values of θ and of v between one-half and unity were posited.
17. See, especially, Atkinson (1970) and Blackorby and Donaldson (1978).
18. If $\varepsilon = 1$, $N_{\varepsilon} = 1 - \Pi_i \, (y_i/m)^{1/n}$.
19. This changes little if familiar differences between urban and rural households are held constant in computing the rural–urban income disparity. Thus, holding constant indicators of household size and structure, the age of the household head, whether the household head is a Communist Party member, and whether the household head is an ethnic minority results in mean rural household income being 41 percent of urban household income. See Khor and Pencavel (2005).
20. Using a maximum likelihood method to compute an entire distribution from grouped summary information, Wu and Perloff (2005) calculate Gini coefficients of household income of 0.338 among rural households and 0.221 among urban households in 1995, values that are somewhat lower than those in Table 3 but the magnitude of the rural–urban difference is similar to the gap we compute. The indicators of income inequality in 1995 among rural households in China in Benjamin et al. (2005) are slightly lower than those in Table 3. For instance, the Gini coefficient for per capita household income in Table 3 for rural Chinese households is 0.358 which is a little larger than the 0.33 reported by Benjamin, Brandt, and Giles for their sample of rural households.
21. The one exception to this statement is the CHNS figure for the ratio of incomes at the 90th percentile to incomes at the 10th percentile among rural households, which is slightly higher in China than the corresponding figure for the United States.

22. To ensure an equal number of households in each quintile, if households at the quintile cutoffs have the same income, they are allocated randomly to the adjacent quintiles.

23. A maximum likelihood test of the symmetry of these transition matrices involves calculating the statistic $\Lambda = \Sigma_{i>j}(p_{ij}-p_{ji})^2/(p_{ij}+p_{ji})$ which has a χ^2 distribution with q $(q-1)/2$ degrees of freedom (with q equal to the number of quantiles). For the transition matrices in Tables 4 through 7, the symmetry hypothesis cannot be rejected with a very high level of confidence (i.e., calculated p values close to unity). See Bishop, Fienberg, and Holland (1975, pp. 282–283).

24. The average quintile move is defined as

$$\frac{1}{5}\left\{\sum_{j=1}^{5}\sum_{k=1}^{5}(|j-k|)p_{jk}\right\}$$

The fraction that remain in the same quintile is defined as $(5)^{-1}\Sigma_{j=1,\ldots 5}\,(p_{jj})$. The immobility ratio resembles Shorrocks' (1978) indicator: $(q-T)/(q-1)$ where T is the trace of the matrix and q the number of quantiles (here 5). As a reference point, if every entry in the transition matrix (i.e., if every value for p_{jk}) were one-fifth (sometimes described as "perfect mobility"), the average quintile move would take the value of 1.6, the immobility ratio would be 0.20, and the adjusted immobility ratio would be 0.52. At the other extreme, if the transition matrix were an identity matrix with unit values on the main diagonal and zeros elsewhere (sometimes described as "complete immobility"), the average quintile move would be 0 and the immobility ratio and the adjusted immobility ratio would each be 1. Evidently, the range of values of the average quintile move is from 1.6 to 0, that of the immobility ratio from 0.20 to 1, and that of the adjusted immobility ratio from 0.52 to 1. Higher values of the average quintile move indicate greater mobility and higher values of the immobility ratio and the adjusted immobility ratio indicate less mobility.

25. Thus, in the income transition matrix in which each element is defined by $\{j, k\}$ where j denotes the income quintile in the initial year and k the income quintile in the final year, $z_i = 1$ if household i occupies an element where $j>k$, $z_i = 2$ if household i occupies an element where $j = k$, and $z_i = 3$ if household i occupies an element where $j<k$.

26. Age is measured in the year 1995 for CHIP, the year 1997 for CHNS, and the year 1998 for the PSID.

27. Estimated standard errors are in parentheses. For continuous variables, marginal effects are partial derivatives while, for discrete variables, the effects report the consequences of a change in the value of the dummy variable from zero to unity. These effects are evaluated at the mean values of the right-hand side variables. "Age" measures years of age of the head of household. "Household size" is the total number of adults and children in the household. "Woman" takes the value of unity for a household headed by a woman. "Communist Party" takes the value of unity for a household head who is a member of the Communist Party. "Minority" takes the value of unity for a household head who reports being an ethnic minority. "Years of schooling" denotes the years of schooling of the household head.

28. The effects are estimated more precisely in CHIP than in CHNS so the statements in this paragraph hold with more confidence for CHIP than for CHNS.

29. The entries in Table 11 and the summary indicators of mobility for China in Table 12 are based on unweighted data. If these Chinese households are weighted by population across provinces, the resulting values are similar. For instance, for CHIP, the value of the average quintile move for households weighted by provincial population are 0.591 for per equivalent adult household income.

30. By stationary, we mean it has the same mean and standard deviation. The assumption of a constant standard deviation, σ, is not egregiously at variance with these data. For instance, for CHIP's total household income, among urban Chinese households, σ in 1993 is 1.10 of σ in 1991 and σ in 1995 is 1.14 of σ in 1991. For total household income, among urban American households, σ in 1996 equals σ in 1994 and σ in 1998 is 1.12 of σ in 1994.

31. For China, ρ_{rs} is the correlation between incomes in 1991 and 1993, ρ_{st} the correlation between incomes in 1993 and 1995, and ρ_{rt} the correlation between incomes in 1991 and 1995. For the United States, ρ_{rs} is the correlation between incomes in 1994 and 1996, ρ_{st} the correlation between incomes in 1996 and 1998, and ρ_{rt} the correlation between incomes in 1994 and 1998.

32. When using incomes averaged over the three years 1991, 1993, and 1995, mean rural household incomes are 46.7 percent of mean urban household incomes. Using 1995 incomes alone, mean rural household income is 46.0 percent of mean urban household income. So at the mean, the rural–urban income gap is almost the same whether using a single year's income or three years' average income.

33. The inequality indicators do not all yield the same rankings between the United States and China, but their general tendency supports the statement in the text.

34. The increases in per capita household income and in per equivalent adult household income were greater than in household income unadjusted for changes in household size and composition. Khan and Riskin (2001) report an annual growth rate between 1988 and 1995 of real per capita household income of 4.48 percent among urban households (somewhat lower than our value of 5.91) and of 4.71 percent among rural households (which is higher than our value of 2.86 percent). As has been emphasized, the sample of households in our empirical work in 1988 and 1995 differs from Khan and Riskin's sample so a difference between their estimates and ours is not surprising. Also, we do not use the same price deflators. Khan and Riskin's growth rate of per capita household income of rural and urban households together is 5.05 percent compared with ours of 5.64 percent.

35. For their sample of households, Khan and Riskin (2001) report an increase in the Gini coefficient for per capita household income of from 0.338 in 1988 to 0.416 in 1995 among rural households and from 0.233 in 1988 to 0.332 in 1995 for urban households. The Gini coefficients of household income in 1988 in Wu and Perloff (2005) are 0.300 among rural households and 0.201 among urban households, values close to those in Table 21.

36. Some intuition for ε may be gained by forming from Eq. (4) the ratio of the marginal social welfare of an increase in household j's income to the marginal social welfare of an increase in household k's income:

$$\frac{\partial V/\partial y_j}{\partial V/\partial y_k} \equiv \Delta_{jk} = \left(\frac{y_k}{y_j}\right)^{\varepsilon}$$

Suppose household k has twice the income of household j. Then giving an extra dollar to household j raises social welfare by 2^{ε} times as much as giving an extra dollar to household k. With $y_k/y_j = 2$, then $\Delta_{jk} = 4$ if $\varepsilon = 2$; $\Delta_{jk} = 32$ if $\varepsilon = 5$; and $\Delta_{jk} = 1{,}024$ if $\varepsilon = 10$.

37. As V is an ordinal representation of preferences, there are no observational consequences from multiplying V in Eq. (5) by $(1-\varepsilon)$ and raising the result to the power of $1/(1-\varepsilon)$ in which case V is linearly homogeneous in m and $(1-N_{\varepsilon})$. When $\varepsilon = 1$, $V = (1-N_{\varepsilon})m$ where N_{ε} for the case where $\varepsilon = 1$ has been defined in the footnote beneath Eq. (1).

38. This involves calculating the value of ε that satisfies $(1-N_{\varepsilon s})/(1-N_{\varepsilon t}) = m_t/m_s$.

39. The growth rates in Table 16 using our trimmed data are similar to those reported by the US Census Bureau. See http://www.census.gov/hhes/www/income/histinc/inchhtoc.html

40. It can be shown that, with $\varepsilon = 2$, N_{ε} is insensitive to increases in income above the median so that a reduction between 1988 and 1995 in the value of N_{ε} when $\varepsilon = 2$ indicates changes in the income distribution in the bottom half of the distribution.

41. One curious feature of the 1995 survey is the high rate of reported female-headed households in urban areas: whereas the fraction of female-headed households in rural areas in 1995 is 4.14 percent and those in urban areas in 1988 is 5.40 percent, the fraction of female-headed households in 1995 in urban areas is 34.00 percent.

42. http://psidonline.isr.umich.edu/

43. No survey of incomes was conducted in 1998.

ACKNOWLEDGMENT

This research was supported by a grant from the Smith Richardson Foundation through the Stanford Institute for Economic Policy Research.

REFERENCES

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, *2*(2), 244–263.

Benjamin, D., Brandt, L., & Giles, J. (2005). The evolution of income inequality in rural China. *Economic Development and Cultural Change*, *53*(4), 769–824.

Benjamin, D., Brandt, L., Giles, J., & Sangui, W. (2008). Income inequality during China's economic transition. In: L. Brandt & T. G. Rawski (Eds), *China's great economic transformation* (pp. 729–775). New York: Cambridge University Press.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge, MA: MIT Press.

Blackorby, C., & Donaldson, D. (1978). Measures of relative equality and their meaning in terms of social welfare. *Journal of Economic Theory*, *18*, 59–80.

Bound, J., Brown, C., & Mathiowitz, N. (2001). Measurement error in survey data. In: J. J. Heckman & E. Leamer (Eds), *Handbook in econometrics (Chapter 59)* (pp. 3707–3745). Amsterdam: Elsevier Science B.V.

Card, D., Lemieux, T., & Riddell, W. C. (2004). Unions and wage inequality. *Journal of Labor Research*, *25*(4), 519–562.

Cowell, F. A., Litchfield, J. A., & Mercader-Prats, M. (1999). *Income inequality comparisons with dirty data: The UK and Spain during the 1980s*. London School of Economics Discussion Paper no. DARP 45, June 1999.

Démurger, S., Fournier, M., & Li, S. (2005). *Urban income inequality in China revisited, 1988–2002*. Unpublished paper, April 2005.

Deng, Q., & Gustafsson, B. (2006). *China's lesser known migrants*. IZA Discussion Paper no. 2152, May 2006.

Fields, G. S., & Zhang, S. (2007). *Income mobility in China: Main questions, existing evidence, and proposed studies*. Mimeograph, Cornell University, Ithaca, NY, December 2007.

Friedman, M. (1962). *Capitalism and freedom*. Chicago, IL: University of Chicago Press.

Glaeser, E. L., & Mare, D. C. (2001). Cities and skills. *Journal of Labor Economics*, *19*(2), 316–342.

Gottschalk, P., & Huynh, M. (2006). *Are earnings inequality and mobility overstated? The impact of non-classical measurement error*. IZA Discussion Paper no. 2327, September 2006.

Griffin, K., & Zhao, R. (1993). *Chinese household income project, 1988* [computer file], Hunter College Academic Computing Services [producer], New York, NY 1992. Inter-university Consortium for Political and Social Research [distributor], Ann Arbor, MI.

Hyslop, D. R., & Imbens, G. W. (2001). Bias from classical and other forms of measurement error. *Journal of Business and Economic Statistics*, *19*(4), 475–481.

Khan, A. R. (2004). *Growth and distribution of household income in China between 1995 and 2002*. Unpublished paper, March 2004.

Khan, A. R., & Riskin, C. (1998). Income and inequality in China: Composition, distribution, and growth of household income, 1988 to 1995. *The China Quarterly*, *154*(June), 221–253.

Khan, A. R., & Riskin, C. (2001). *Inequality and poverty in China in the age of globalization*. New York: Oxford University Press.

Khor, N., & Pencavel, J. (2005). Income disparities and income mobility in China. Paper prepared for the conference on "China's Policy Reforms: Progress and Challenges", Stanford Center for International Development, Stanford University, October 2005.

Khor, N., & Pencavel, J. (2006). *Household income inequality, income mobility, and labor supply in China and the United States*. Unpublished paper, February 2006.

Riskin, C., Zhao, R., & Li, S. (2000). *Chinese household income project, 1995* [computer file], ICPSR version, Amherst, MA, University of Massachusetts, Political Economy Research Institute [producer]. Inter-university Consortium for Political and Social Research [distributor], Ann Arbor, MI, November 2000.

Shorrocks, A. F. (1978). The measurement of mobility. *Econometrica*, *46*(5), 1013–1024.

Wu, X., & Perloff, J. M. (2005). *China's income distribution, 1985–2001*. Unpublished paper, February 2005.

Ximing, Y., Sicular, T., Shi, L., & Gustafsson, B. (2008). Explaining incomes and inequality in China. In: B. A. Gustafsson, L. Shi & T. Sicular (Eds), *Inequality and public policy in China* (pp. 88–117). New York: Cambridge University Press.

APPENDIX

Chinese Data

The Chinese Household Income Project 1988 and 1995.

Both sets of data are publicly available through the Inter-university Consortium for Political and Social Research and more fully described in the relevant codebooks. This appendix describes the construction of the pertinent data for this project. The data for both rounds of the survey are available in four files. Here we report the initial number of observations from the raw data:

Data Files	1988	1995
Rural individual files	51,352	34,739
Rural household files	10,258	7,998
Urban individual files	31,287	21,698
Urban household files	9,009	6,931

1995 Survey

For urban households, total income is constructed by summing over the total annual income of all members in 1995 (variable a51). For rural households, this total is reported directly (q600). Income is the sum of labor income, property income, transfer income (including retirement income), and ''income from household sideline production'' and household income is the sum of reported income from all family members. The largest component of total income is labor income. To address outliers, the data were then trimmed by dropping the top 0.5 percent and the bottom 0.5 percent of households in rural and urban areas respectively (148 observations in all). In addition, observations for which household heads are younger than 20 years are omitted (16 observations). The resulting sample consists of 6,863 urban households and 7,917 rural households.[41]

To minimize measurement error, a large amount of time was devoted to the verification of the income responses in 1995 and in years prior to 1995. This was tackled by inspecting the run of incomes over the years for each household and marking abrupt or peculiar values. Sometimes these households were dropped from the sample. In other instances, it seemed plausible to modify the recorded entry on income. This would occur, for example, when zeros were missing in a particular year.

In the analysis reported in the body of the chapter, the definition of income follows that constructed by the NBS. However, we investigated the consequences of other definitions. For instance, Khan and Riskin's (2001) income construct approximates disposable household income and incorporates government and other in-kind transfers. For rural households, this means the inclusion of wages, pensions, income from rural enterprises and farming, and net transfers from collectives and the state. For urban households, disposable household income consists of wages, pensions, the income of nonworkers, property income, and net public transfers and subsidies. The table below presents some descriptive statistics on alternative definitions of household income in 1995. The NBS column describes household income as used in the analysis of this chapter while the column listed ALT approximates to the income concept used by Khan and Riskin. Naturally, the central tendency of household income when using Khan and Riskin's definition is higher than that used by the NBS. However, the key qualitative differences – the lower incomes and greater income dispersion in rural than in urban areas – hold for both definitions of income.

	Total Household Income		Per Capita Household Income		Per Adult Equivalent Household Income	
	NBS	ALT	NBS	ALT	NBS	ALT
Urban and rural						
Mean	10,204.3	13,406.5	3,122.9	4,062.7	3,883.8	5,058.3
Median	8,854.0	10,683.7	2,556.6	3,090.4	3,267.8	3,946.2
SD	6,881.2	18,245.2	2,400.1	4,939.1	2,840.9	6,242.9
Gini	0.355	0.386	0.403	0.383	0.403	0.380
Urban						
Mean	13,741.0	17,287.1	4,572.0	5,722.9	5,599.2	7,013.8
Median	12,364.0	14,467.0	4,052.5	4,736.0	5,024.1	5,873.9
SD	6,796.7	22,885.4	2,321.5	5,919.0	2,712.9	7,577.2
Gini	0.257	0.311	0.265	0.316	0.265	0.316
Rural						
Mean	6,326.1	9,151.0	1,533.9	2,242.2	2,002.6	2,913.8
Median	5,272.0	6,828.5	1,242.8	1,626.8	1,646.7	2,148.3
SD	4,457.0	9,435.1	1,157.6	2,528.4	1,443.9	3,156.8
Gini	0.354	0.395	0.358	0.407	0.358	0.407

1988 Survey

The construction of the household files for 1988 is slightly more complicated than the 1995 data. Out of the 31,759 observations in urban areas, the responses to the question about self-reported relationship indicate 9,021 head of households and yet a total of 9,009 unique households are identified. Among the urban households, 27 report more than one head of household. Also these households report more than one spouse to the head of household (23 observations). In such cases, the oldest member was selected as the head of household and the reported spouse closest in age was chosen to be the spouse. This seems the sensible procedure as, for some households, each member is coded as the household head including children as young as one year in age.

For heads of households for whom demographic and education variables are missing or those households with no head of households at all, the missing values are replaced with the reported values of the spouses. As is the case for the 1995 data, the 1988 data were adjusted by excluding those households with household heads younger than 20 years (83 observations) and then trimmed by excluding those households in the top and those in the bottom 0.5 percent of household income. This yields a sample of 18,947 households, 10,080 in rural areas and 8,867 in urban areas.

The China Health and Nutrition Survey

This is an ongoing international collaborative project between the Carolina Population Center at UNC at Chapel Hill, the National Institute of Nutrition and Food Safety, and the Chinese Center for Disease Control and Prevention. It covers Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong. As with the other data, the values of adult equivalent household income were trimmed to remove the bottom and the top 0.5 percent of observations.

To construct a consumer price index, the CHNS research team used a consumer goods basket specified by the government and published urban price data to create the urban cost of this basket. The basket includes 57 items of goods. Urban prices in each province come from the NBS volumes. Then CHNS urban and rural price data are used to create a ratio of urban and rural costs for elements of this consumer goods basket. In this way, they construct the yuan cost of this basket for each time period for urban and rural areas in each province in the CHNS. They set China food

costs for urban Liaoning province for 1988 equal to 1.0 and all other prices are indexed to this.

DATA FOR THE UNITED STATES

Current Population Survey

From the 1996 CPS March Demographic Survey, individuals are included in all types of households (both civilian and military) except those individuals living in group quarters. Excluding those households containing these individuals yields a total of 56,873 households. Using the geographical indicators for the household, the urban–rural distinction is drawn on the basis of whether the household lives in a metropolitan area. (This information is not provided for 1,108 households.) In constructing household level variables, only those defined as relatives and unmarried partners of the reference individual were included (i.e., nonrelatives, housemates, and boarders were excluded). In addition, only households are included where the reference individual is at least 20 years old. This results in losing 283 households. These restrictions yield 142,606 individual person records and 55,766 household records. Household income is formed by adding reported income in all categories for each individual in the household. These income categories include wage and salary, earnings, interests, and dividends in addition to governmental transfers such as unemployment compensation and social security benefits. Top-coding of some income components affected 2.15 percent of the households with most of these cases attributable to top-coding of earnings. Nothing has been done to adjust for top-coding. Finally, we trim the sample of households by dropping the bottom 0.5 percent and top 0.5 percent of household income. This results in a final sample of 54,770 households.

Panel Study of Income Dynamics

The PSID data come from the online data center.[42] The characteristics of these US data were limited to conform to those applied to the Chinese data. Owing to changes in reported income variables, our sample for the PSID includes the later survey years of 1994–1999.[43] Initially, this includes 38,141 individuals per year. The sample size was reduced because of a number

of changes made to the PSID in 1997. Thus, the number of heads of households fell from 10,972 in 1994 to 7,176 in 1999.

The income variable we use is from the Income Plus files, which contain data on family income and its components, notably the labor earnings of the head and spouse. Because we seek to construct a balanced panel on income, we restrict the sample of the individuals to those who are always in the sample and who are either heads of households or their respective spouses. Eliminating the over-sampled population further reduces the sample size. Finally, the data were trimmed by dropping the top and bottom 0.5 percent of incomes.

WHY ARE JOBS DESIGNED THE WAY THEY ARE?☆

Michael Gibbs, Alec Levenson and Cindy Zoghi

ABSTRACT

In this chapter we study job design. Do organizations plan precisely how the job is to be done ex ante, or ask workers to determine the process as they go? We first model this decision and predict complementarity among these following job attributes: multitasking, discretion, skills, and interdependence of tasks. We argue that characteristics of the firm and industry (e.g., product and technology, organizational change) can explain observed patterns and trends in job design. We then use novel data on these job attributes to examine these issues. As predicted, job designs tend to be "coherent" across these attributes within the same job.

☆The data used in this paper are restricted-use; we thank Brooks Pierce for his guidance in analyzing them. We thank John Abowd, Gary Becker, John Boudreau, Susan Cohen, Jed DeVaro, Alfonso Flores-Lagunes, Kathryn Ierulli, Ed Lawler, Canice Prendergast, and workshop participants at the American Economic Association Annual Meeting, Aarhus School of Business, BLS, Cornell, Illinois, LSE, the NBER Summer Institute, the Society of Labor Economists, Universidad Carlos III de Madrid, and USC for their comments. Michael Gibbs gratefully acknowledges the hospitality of the Center for Corporate Performance at the Aarhus School of Business, and funding from the George Stigler Center for the Study of the Economy and the State, and the Otto Moensted Foundation. All views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the US Bureau of Labor Statistics.

Jobs, Training and Worker Well-Being
Research in Labor Economics, Volume 30, 107–154
Copyright © 2010 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 0147-9121/doi:10.1108/S0147-9121(2010)0000030007

Job designs also tend to follow similar patterns across jobs in the same firm, and especially in the same establishment: when one job is optimized ex ante, others are more likely to be also. There is evidence that firms segregate different types of job designs across different establishments. At the industry level, both computer usage and R&D spending are related to job design decisions.

1. INTRODUCTION

Job design is a fundamental issue in organization design. Which tasks should be put together in the same job, what skills and training are needed, what decisions the employee is allowed to make, with whom the employee works, and related questions are crucial for efficiency and innovation. These issues have long been a focus of social psychology, which has a large literature on effects of job "enrichment" on intrinsic motivation. By contrast, job design has been underemphasized in economics, with some notable exceptions such as Adam Smith's (1776) discussion of specialization.

Empirical evidence suggests that there are patterns and trends in job design. For example, the management research literature and evidence from large organizations (Cohen & Bailey, 1997; Lawler, Mohrman, & Benson, 2001) suggest a trend in recent decades toward teams and human resource practices associated with job "enrichment," i.e., multitasking instead of specialization, and greater employee discretion. In addition, this job design approach seems to be positively associated with organizational change (Milgrom & Roberts, 1990, 1995; Caroli & Van Reenen, 2001). Finally, a substantial literature argues that organizational change in recent years has been skill-biased, leading to increasing returns to skills and a greater emphasis on higher-skilled workers in firms that have undergone change (Autor, Katz, & Krueger, 1998; Bresnahan, Brynjolfsson, & Hitt, 2002; Autor, Levy, & Murnane, 2003; Zoghi & Pabilonia, 2004).

In this chapter we present an economic analysis of job design. First, we present a simple model of inter-task learning that can provide an explanation of trends toward broader job design and greater worker discretion, and the association of job design attributes with organizational change. The model is based on a straightforward idea: combining interdependent tasks in a job may enable the worker to learn process improvements. If this effect dominates gains from specialization, then multitasking leads to greater productivity. Learning should be greater for high-skill workers who

are given discretion. Thus, interdependence may lead to multitask jobs, and greater discretion and skills. We then argue that job design should be related to characteristics of the firm's environment – its product, industry, and technology – yielding economy-wide patterns of job design within firms, and within establishments in the same firm.

The predictions about economy-wide patterns of firm characteristics and job design are relatively new to both the economic and social psychology literatures on job design. The empirical literatures have previously ignored such patterns because the existing data are not drawn from representative national samples. Lacking data with which to test such predictions, the theoretical literatures similarly have not explored them in depth. One exception from the theoretical literature is Morita (2001), which focuses only on specificity of human capital and not other aspects of job design. Thus a contribution of the chapter is the job design predictions at an economy-wide level.

The second part of the chapter analyzes a unique dataset that provides the first nationally representative view of the distribution of job design characteristics. The Bureau of Labor Statistics (BLS) National Compensation Survey (NCS) measures job design attributes, including multitasking, discretion, skills, and interdependence. As predicted, we find that all four are strongly positively correlated. At the job level, there is a strong tendency toward "coherent" job design, meaning that jobs tend to be high, medium, or low on all four attributes, relative to the occupation median for each attribute. At the establishment level, there is a tendency for firms to choose either a "modern" approach (many jobs high on all design dimensions) or a "classical" approach (many jobs low on all dimensions). This is consistent with our arguments that job design approaches vary with the firm's product and market characteristics. At the firm level, there is a tendency to push job design toward extremes, choosing modern design in some establishments and classical design in others. This is consistent with multi-establishment firms using establishments to isolate modern and classical jobs from each other to maximize the benefits of job design. At the industry level, both R&D spending and computer usage are associated with modern job design.

2. A SIMPLE MODEL OF MULTITASKING, INTERDEPENDENCE, AND DISCRETION

We now present a simple model of job design based on Lindbeck and Snower (2000) and Gibbs and Levenson (2002). We augment the Lindbeck

and Snower approach by considering employee discretion. Our first results are similar to the previous literature (Milgrom & Roberts, 1990; Holmstrom & Milgrom, 1991, 1994; Morita, 2001; Dessein & Santos, 2006) in providing an argument for complementarity of specific job design components. We then discuss implications for the distribution of job design characteristics within establishments compared to the firm as a whole, and at the economy-wide level. After the model, we discuss several related empirical predictions suggested by our approach. The model and other predictions are developed explicitly with the goal of generating testable predictions for the dataset used in this chapter.

Consider a setting where a firm has to allocate production between two workers. It has the choice of specializing jobs, or of using multitasking (where workers work independently from each other, producing the entire product or service themselves). In the case of multitasking, it also has the choice of deciding how workers should allocate their time between tasks, or giving them discretion to decide this for themselves. Our analysis is intended to shed light on factors that might tip the balance of job design toward specialization or multitasking, and toward centralization or decentralization. For this reason, we do not model some related issues. In particular, our analysis understates the advantages of specialization, because we force the ratio of specialized workers to be one-to-one. Allowing firms to deploy different ratios of workers to each task, or to have some multitask and some specialized workers, would improve the firm's ability to exploit differences in productivity across the two tasks. Similarly, we do not model agency problems ensuing from worker discretion. That is partly because we do not have good incentive variables in our dataset, and also because our focus is on how the job design itself affects productivity independent of any incentive effects.

Consider a firm with two workers, each with one unit of time to perform assigned tasks. There are two possible methods of production. In one, both workers multitask, producing the entire product via a Cobb–Douglas production function, and total firm output is the sum of individual worker outputs. In the other, both workers specialize, and work from their tasks is combined within the Cobb–Douglas production function to get total output. The production function is $Q = X_1 \times X_2^{\alpha}$. Their marginal product of effort on a task equals s.

Thus, if the workers specialize and their work is combined, output is $Q = s^{1+\alpha}$. As in Becker and Murphy (1992), assume a constant coordination cost C if workers specialize, but none if they multitask:

$$Q_{\text{specialized}} = s^{1+\alpha} - C \tag{1}$$

Now consider the opposite case, where workers spend some time on each task. The key idea in this chapter is *inter-task learning*: in performing one task, the worker may improve output on the other. For example, a worker who performs both tasks should better understand what to emphasize in performing each task, so that the outputs from both tasks fit together better, leading to lower costs or better quality. Exposing a worker to a broader set of tasks also may lead to more innovation and creativity. Using the familiar example of academia, most universities are organized to combine teaching and research, because in most cases working on one improves work on the other. Similarly, interdisciplinary research is often encouraged because it tends to lead to more creative new research topics.

Define τ as the fraction of time that a multitasking worker spends on task 1, with $1-\tau$ for task 2. To capture inter-task learning, which is only relevant for multitasking workers, the extent that output improves on a task is proportional to time spent on the other task:

$$X_1 = s\tau + k(1-\tau); \quad X_2 = s(1-\tau) + k\tau$$

where $k =$ the degree of inter-task learning. There are thus two competing effects on worker productivity. One is the standard gains from specialization s, which applies to all workers; the other is the gains from inter-task learning k, which applies only to multitasking workers. We do not assume that one effect is larger than the other. Output for a single multitasking worker i is:

$$Q_i = (s\tau + k(1-\tau))(s(1-\tau) + k\tau)^{\alpha}$$

τ is chosen by the firm to optimize Q_i:

$$\tau^* = \frac{s - \alpha k}{(1+\alpha)(s-k)}; \quad 1-\tau^* = \frac{s\alpha - k}{(1+\alpha)(s-k)} \tag{2}$$

Given the allocation of time between the two tasks, individual worker output is given by substituting τ^* and $1-\tau^*$ into Q_i above. Total output is twice this for two multitasking, independent workers:

$$Q_{\text{multitask}} = 2\alpha^{\alpha}\left(\frac{s+k}{1+\alpha}\right)^{1+\alpha} \tag{3}$$

For example, if $k = 0$ and $\alpha = 1$, then $Q_{\text{multitask}} = 1/2 \times s^2$, and $Q_{\text{specialized}} = s^2 - C$, which is greater than $Q_{\text{multitask}}$ as long as C is not too large. The greater the coordination costs, the more likely is multitasking to be optimal rather than specialization. In Eq. (2), for multitasking with $\tau \in (0,1)$, α cannot be too different from 1 in either direction.

Similarly, comparing Eqs. (1) and (3), as α diverges from 1 in either direction, specialization is more likely to be the best design. Thus we should see multitasking only if comparative advantage is not too strong. The effects of higher marginal product s are also ambiguous, since higher s increases output for both specialized and multitask jobs.

In the appendix, we show that there is always some range of parameter values for which multitasking is more efficient than specialization. Holding s and α fixed, the larger is the opportunity for inter-task learning k, the more likely is this to be the case.

2.1. Multitasking and Interdependence

An immediate result of Eqs. (1) and (3) is that multitask jobs are more likely to be optimal; the more important is inter-task learning:

$$\frac{\partial Q_{\text{multitask}}}{\partial k} > 0, \quad \text{while} \quad \frac{\partial Q_{\text{specialized}}}{\partial k} = 0 \tag{4}$$

In this view, a primary cause of multitasking – which reduces traditional gains from specialization – is that it allows the worker to learn about production and make continuous improvements. The degree of specialization is limited not just by coordination costs (Becker & Murphy, 1992), but also by inter-task learning opportunities.[1] For workers to learn on the job, multitasking is important because task interdependencies are an important source of inefficiencies in production, and one that is exacerbated by specialization. Thus, complex production processes (greater task interdependence) are more likely to use multitask jobs.

Our approach stands in contrast to Morita (2001), who addresses the conditions under which an economy will have an equilibrium with jobs that emphasize continuous process improvement, training, and specific human capital versus an equilibrium with jobs that have general human capital, less training, and little to no continuous process improvement. In Morita's model, workers learn how to perform specialized tasks that have a return only to the firm currently employing them – hence the accumulation of firm-specific human capital. A key issue Morita sought to address was lower turnover (and greater training) in Japan versus the United States. In our model, in contrast, learning how to perform multitask jobs does not lead to the accumulation of firm-specific capital. Moreover, our predictions do not lead to an equilibrium in which all jobs in an economy are either specialized or not specialized, as is the case for Morita's model.

The role of task interdependence in this model is similar to what Milgrom and Roberts (1990) call complementarities among elements of the firm's strategy. In their formulation, complementarities mean that the marginal returns to adopting one element are increasing in the level of the other elements. In their case, they examine aspects such as technology adoption, marketing, and engineering. In their model, if there are complementarities among these, then it makes economic sense for the firm to make coordinated changes among all of them at the same time. For example, introducing computer-aided design technology makes it cheaper for the firm to adapt a broader product line and to update its products more frequently, which is reinforced by an engineering approach that designs production processes more quickly using cross-functional teams, and by a marketing approach that emphasizes lower prices, faster delivery, and smaller batch sizes (more customized product lines). The main difference between Milgrom and Roberts (1990) and our model is the focus: Milgrom and Roberts focus on technology changes and organization design; we focus on job design.

Holmstrom and Milgrom (1991, 1994) more directly consider job design: the firm's decision is over whether to hire the person directly as a regular employee or as an independent contractor. Hiring as a regular employee means greater supervision and less discretion than hiring as an independent contractor. In our approach, as detailed in the next section, discretion and supervision are central to the firm's decision-making process. However, instead of deciding over a relationship as regular employee versus independent contractor (representing two polar opposites of discretion and supervision), in our approach the firm selects different amounts of discretion and supervision for a range of "regular employee" jobs.

2.2. Multitasking and Discretion

Another important job design characteristic is the degree of discretion (decentralization) given to an employee (Ortega, 2004; Zoghi, 2002). When there is learning in a multitask job, discretion allows the worker to test new methods of production to solve problems and implement improvements (Jensen & Wruck, 1994). In our model, a simple way to capture this idea is that discretion allows the worker to adjust the allocation of time τ depending on circumstances. For example, suppose the production environment k (or s, k/s, or α) is stochastic, and ex ante the firm knows the distribution of k but not its specific value. If workers perform both tasks, they observe the state of the world before choosing their allocation of time,

allowing them to observe in real time the relative value of focusing on one task or devoting time to both. If they are specialized, they do not possess this knowledge because they do not perform the second task, and regardless have no time allocation decision to make. If workers are given discretion, they can choose τ based on this knowledge, though at some agency cost D.[2] Otherwise, the firm chooses τ without this knowledge. Using the worker's knowledge can improve output.

$$E[Q_{\text{multitask|discretion}}] \geq E[Q_{\text{multitask|centralization}}] \tag{5}$$

For proof of Eq. (5), see Appendix B. Moreover, discretion will tend to be more valuable in more uncertain production environments. From Eq. (3), Q is convex in s, k, s/k, and α. Therefore, expected output will be higher when variance in any of these parameters can be exploited by the worker. Unfortunately, solving for the optimal time allocation τ^* when production is stochastic does not yield closed form solutions, even for simple cases (e.g., binary k or α). However, combining these ideas and the case in Eq. (4) above, a reasonable prediction to test with our data is that discretion should be complementary with multitasking, especially in more uncertain environments.

We do not model incentives. Giving a worker discretion creates agency costs. The firm would presumably respond by implementing an incentive scheme to better align incentives. Thus, the benefits of discretion would in practice be net of agency costs. In datasets similar to ours but including information on incentives, it would be interesting to study whether incentives are more likely to be used, and are stronger, the greater is the use of discretion, multitasking, and interdependence.

Putting these two arguments together, the model predicts complementarity among multitasking, interdependence, and discretion. It also predicts complementarity among specialization, lack of interdependence, and centralization. This suggests two patterns of job design. The first we will call "classical" job design: specialized jobs with little discretion. The second we will call "modern" job design because it matches the apparent trend: "job enrichment" as described in the behavioral literature, using multitasking and more worker discretion. Both types of jobs should be observed in the economy (or industry, or firm). The extent to which we expect to see one or the other depends on the importance of gains from specialization versus inter-task learning. We expect to see "classical" jobs more where interdependence is lower, and "modern" jobs more where interdependence is higher.

Our model shares some similarities with Lindbeck and Snower (2000). In both cases, multitask learning provides the foundation upon which implications for specialization and job design are derived. Lindbeck and Snower (2000), however, consider the roles of technological change that promotes task complementarities (similar in spirit to Milgrom & Roberts', 1990 complementarities), changes in worker preferences for multitask work, and advances in human capital that makes workers better able to multitask; they do not consider other aspects of job design such as discretion. By addressing discretion and the degree of supervision, we indicate potential additional insights into firms' job design choices.

Our model also shares some similarities with Dessein and Santos (2006), which was developed contemporaneously. Dessein and Santos (2006) address the relationships among specialization, discretion, the ease of communication between employees about tasks and their outcomes, and uncertainty in the economic environment. Their goal, similar to ours, is to provide a model that can explain organizations' decisions to create modern versus classical jobs. Their emphasis on uncertainty provides similar predictions as our focus on product complexity, however, with important differences: their approach is better suited for exploring how the external economic environment influences firms' job design decisions; our approach focuses more on how product characteristics influence job design decisions. Moreover, they do not address the role of skills, which we do in the next section.

2.3. The Role of Skills

Skills play a central role in labor economics research, so it is of interest to consider their role in this context. There are two general off-setting effects. The first is that gains from specialization may be complementary to skills. For example, specialization may increase returns on investments in skills in two ways (see Murphy, 1986). First, specialization of training may lower training costs if there are fixed costs to learning new topics. Second, focused work may lead to economies of scale in skill acquisition on the job. For these reasons, we might see more highly skilled workers in more specialized jobs.

A countervailing effect is that skills may facilitate on-the-job learning. If more highly skilled workers are better able to learn on the job, then skills will be complementary to discretion. Returns to skills would be higher in more complex work environments, where the scope for inter-task learning is higher. This effect is suggested by the literature on skill-biased technical

change. Much of that literature (Autor et al., 1998, 2003; Goldin & Katz, 1998) has focused on the relationship between technology change and wages, but job design considerations are also important (Autor, Levy, & Murnane, 2002). If certain types of technological change complement problem-solving or abstract-thinking skills (Levy & Murnane, 2005), they may increase the strength of inter-task learning.

Which effect dominates is an empirical question. If skills are more complementary to specialization, then we should see more highly skilled workers given narrow jobs with low discretion – to became masters of their specialized trades. If skills are more complementary to discretion and multitasking, then we should see more highly skilled workers given more enriched jobs.

As a prelude to the empirical work below, it is worth noting that by "skills" we mean the ability to perform the tasks that are needed for a job. Because tasks differ in the skills needed to execute them, we do not assume that what defines "highly skilled" for one set of jobs or for an occupation is the same as what defines skills in another set of jobs or occupation. In particular, we are concerned about skills that are more specific than can be described by total years of schooling, general degree attainment (i.e., high school graduate versus undergraduate degree versus graduate degree), and total years of labor market experience. Though not exactly the same, previous measures of occupation-specific experience (Shaw, 1987; Neal, 1999) are the closest analogy from the existing literature to our concept of the skills needed to perform specific job tasks. As the discussion below details, the job-based data we analyze contain a more precise measure of task- and job-relevant skills than standard employee-based datasets.

2.4. The Role of Product and Process Characteristics

Our argument is that a primary reason for multitasking is to facilitate continuous improvement by workers as they perform their jobs. An alternative way for the firm to choose effective production methods is to invest in ex ante optimization. In fact, an important influence on the early job design literature and practice is industrial engineering, a formal method for ex ante optimization pioneered by Frederick Taylor ("Taylorism") and others in the early 20th century. Ex ante optimization should tip the balance away from multitasking and toward specialization, since it implies that there will be less scope for workers to learn improvements on the job.

This helps provide additional predictions about patterns of job design within establishments, firms, and industries.

Consider ex ante optimization of production methods as an investment by the firm. Our model might be extended to allow the firm to invest in ex ante process improvements at some cost. This would increase s and/or α, but reduce opportunities k for workers to make continuous improvements. A greater investment in better methods should therefore induce more use of classical job design. The expected return on investments in ex ante optimization depends on the degree to which it uncovers methods close to the optimum, and the extent to which the efficiency gains are expected to be reaped in the future. These depend on the complexity, predictability, and stability of the firm's product and environment.

First consider product or process complexity. Greater complexity (e.g., more parts; modules in a software program; broader product line) should imply greater cost to ex ante perfection of production methods. The cost of optimizing the manufacture of a tin can (less than half a dozen parts) is substantially lower than optimizing the manufacture of a diesel engine (2,000 or more parts). Moreover, in the diesel engine, the parts have to work together well – there is high interdependency. Such interdependencies tend to be the kind of situations where ex ante optimization is more difficult, quality problems arise, etc.

A second important characteristic of the product or process is the extent to which it is unpredictable. Consider management consulting. Each client engagement is typically different from the last. Some processes and methods can be reapplied, but new methods or applications often need to be developed. Moreover, judgment as to what methods to apply may be required. To the extent that situations arise over and over, the consulting firm may be able to develop standard methods and provide employees with a menu of choices from which to select. However, if any of the work is idiosyncratic and unforeseeable, some optimization will have to occur in real time.

A third important product or process characteristic is stability. This plays out both backward and forward in time. The longer a product has been produced with few or no changes, the more is known about how to make it efficiently, and the lower is the potential for inter-task learning. The longer the firm expects to make the same product in the future, the greater the expected returns on ex ante optimization, leading to greater investments in ex ante optimization.

These factors (complexity, predictability, and stability) influence the return on investments in ex ante optimization of methods, and therefore optimal job design. If the return is small, the firm will invest less in ex ante

optimization, and there are greater possibilities for employees to engage in continuous improvement. Continuous improvement is more likely to be successful with a modern approach to job design, and vice versa. Therefore, for groups of workers producing products or using processes that have similar complexity, predictability, and stability, job design should be similar. The more similar these factors for two workers, the more would we expect their job designs to be similar to each other in terms of multitasking or specialization; discretion or decentralization; and degree of skills. This should even apply across jobs that are in different occupations.

This leads to several useful empirical predictions. First, firms should tend toward choosing a similar job design approach (on the spectrum from classical to modern job design) for all jobs within the same firm. This is consistent with Milgrom and Roberts' (1990) complementarities model and with research on the effects of adoption and use of "high performance work systems" on productivity and profitability of organizations (Appelbaum & Batt, 1994; Cappelli & Neumark, 2001; Ichniowski & Shaw, 1995; Ichniowski, Shaw, & Prennushi, 1997; MacDuffie, 1995). Many of these studies find that while the adoption of a single policy does not affect measurable outcomes, there are complementarities between policies that can have real effects.

The complementarities should even apply to workers in different occupations. For example, if a firm gives its production workers greater discretion and more tasks than is typical, we predict that the same firm is more likely to also give its secretaries greater discretion and more tasks. Thus we expect a clustering of high levels of multitasking, discretion, skills, and interdependence within some firms, medium levels at other firms, and low levels at still other firms. In social psychology, Porter, Lawler, and Hackman (1975) make a similar conjecture, which they do not test. Note though that high, medium, and low are relative terms. The prediction is about multitasking, etc. *relative to their occupational norms.*

Note that we do not conclude that modern jobs are optimal for all establishments. If a firm employs multiple strategies across its product line, segmenting the strategies by establishment may be a preferred way of accomplishing its objectives. For example, consider a large firm with a diversified product line spanning both high and low margin products, such as General Electric (GE). GE separates into different divisions (and establishments) the design, engineering, marketing, and production of light bulbs versus jet engines. Though the benefits of modern job design may accrue to both types of production, they should exhibit a greater rate of return in jet engines where the degree of complexity is much greater than it is

for light bulbs. Thus an optimal job design strategy may include adopting different degrees of modern job design across establishments making different products or servicing different customer segments.

Such patterns should be stronger within establishments than within firms as a whole, given differences in the degree of product diversification across firms. At a naïve level, product attributes are likely to be more similar within than across establishments because of product diversity within firms. Less naïvely, establishments are groupings of employees *chosen by the firm*. Because workers are grouped together by choice, it is more likely that the products, customers, technology, etc. that they work with are the same as their colleagues' in the same establishment, compared to employees randomly chosen from the same firm but different establishments. Moreover, if workers are put together at a site when their work is highly interdependent, establishments can in a sense be viewed as teams. If their work is interdependent, then it is even more likely that product and technology attributes will affect them similarly.

Finally, this general prediction should also apply, though more weakly, within industries. Within an industry, products and processes should be more similar than in the economy as a whole. This implies that the returns to investments in ex ante optimization should vary by industry, and there should be patterns of ex ante optimization or continuous improvement across industries. Therefore, industries should show some tendency toward greater use of modern or classical job design approaches.

This logic might also help explain a recent trend toward "modern" jobs (Caroli & Van Reenen, 2001). The past few decades have exhibited rapid change, due to modern manufacturing and flexible production methods, information technology and technological change, shorter product cycles, and increasing emphasis on customization and complex product lines (Milgrom & Roberts, 1990, 1995). All reduce the returns from investing in industrial engineering, and increase the returns to continuous improvement. In a changing environment, there is greater scope for workers to develop improvements and aid implementation of change, because old methods are less likely to be optimal. We now turn to a description of the data that we employ to test these ideas.

3. DATA

Our empirical analyses use a novel dataset that contains information on job design from a nationally representative sample of establishments in the

United States. The NCS is a restricted-use dataset collected by the BLS. It covers the nonagricultural, nonfederal sectors of the US economy. Our data are from 1999. The data were collected by field economists who visited sampled establishments and randomly selected 5–20 workers from the site's personnel records, depending on establishment size. Through interviews with human resources representatives, detailed information about the jobs those workers hold was obtained.

The data include information on occupation and union status of each job, industry, whether the establishment is privately owned or public (state or local government), earnings, and an indicator for use of incentive pay.[3] No demographic information about the worker is collected. The most unusual feature of the dataset is the "leveling factors," which are intended to measure various job design attributes consistently across occupations. These factors are based on the federal government's Factor Evaluation System, which is used to set federal pay scales.[4] There are 10 different leveling factors, or job design attributes, of which we use 5 in this chapter[5]: knowledge; supervision received; guidelines; complexity; and scope & effect. Here we provide a brief synopsis of each and how they correspond to the concepts from our theoretical discussion. All are measured on Likert scales with ranges varying from 1–3 to 1–9.

1. *Knowledge:* This measures the nature and extent of applied information that the workers are required to possess to do acceptable work – this is quite similar to the general notion of human capital, though it differs substantially from the typical operationalization used by labor economists (measuring education/years of schooling and years of general labor market experience). Values of 1–2 roughly correspond to skills required to do simple, routine, or repetitive tasks; 3 is the level of skills required to do standard clerical assignments, resolve recurring problems, or operate and adjust varied equipment for purposes such as performing standardized tests or operations; 4 is at the level of an apprenticeship or someone who can perform nonstandard procedural assignments and resolve a wide range of problems; 5 is at the level of a college graduate who has mastered the basic principles, concepts and methodology of a professional or administrative occupation, and/or who can solve unusually complex problems; and so on. Thus, larger values imply greater knowledge. This factor corresponds quite well to our *skills* job design attribute.
2. *Supervision received:* This measures the nature and extent of supervision and instruction required by the supervisor, the extent of modification and participation permitted by the employee, and the degree of review of

completed work. Larger values correspond to *less* supervision. Values of 1–2 indicate substantial supervisory control with minimal employee input; 3 implies some autonomy for the employee to handle problems and deviations; 4–5 indicate that general objectives are set by the supervisor while the worker has more responsibility for implementation and there is little review of the completed job. This factor corresponds to some dimensions of *discretion* in our discussion above. We use it, along with the next factor, to proxy for that concept.

3. *Guidelines:* Measures how specific and applicable the guidelines are for completing the work, and the extent of judgment needed to apply them. As with supervision received, larger numbers correspond to *less* use of Guidelines. Values of 1–2 signify that detailed guidelines are available that are applicable in most situations that are likely to arise; 3 indicates that, while guidelines are available, the worker must judge whether they are applicable, and how to adapt them; 4–5 indicate that few guidelines are available or applicable to completing this job. Thus, we interpret both supervision received and guidelines as indicators of our concept of *discretion*.[6]
4. *Complexity:* Complexity measures two things: the extent to which the job has multiple dimensions, in terms of the nature, number, variety, and intricacy of tasks or processes; and the extent to which the job has unpredictability, due to the need to assess unusual circumstances, variations in approach, and the presence of incomplete or conflicting data. The former is closer to what we mean by multitasking as the opposite of specialization, though unpredictability also suggests variation in tasks. Moreover, complexity is positively associated with interrelationships between tasks. In our discussion of job enrichment, we argued that an important reason for multitasking is to design jobs so that employees see complex interactions between the most complementary tasks. Thus, the NCS Complexity corresponds reasonably well to our concept of *multitasking*.
5. *Scope and effect:* Scope and effect measures the extent to which the employee's work has impacts on activities and persons in (and beyond) the organization, for example by affecting the design of systems, the operation of other organizations, the development of programs or missions. As scope and effect gets larger, the impacts get larger. This measures the interdependence of a job with other processes and jobs in and beyond the organization, rather than interdependence between tasks within the same job. However, it seems likely that greater interdependence between jobs will be positively correlated with greater

interdependence between tasks within jobs, indicating that overall interdependence is higher. We interpret this as a proxy for *interdependence*.[7]

4. RESULTS

4.1. Bivariate Relationships between Job Characteristics

Table 1 shows the Spearman's rank-order correlations among the five factors. The correlations are high, consistent with our prediction that there should be positive relationships among multitasking, discretion, and interdependence. Table 2 replicates the bivariate relationships from Table 1 using ordered logits, predicting multitasking as a function of either discretion (measured by either guidelines or supervision), skills, or interdependence; guidelines as a function of supervision, skills, or interdependence; supervision as a function of skills or interdependence; and skills as a function of interdependence. Each cell represents a separate regression, with the row naming the dependent variable and the column naming the independent variable. The first number in each cell shows the estimated ordered logit coefficient.

Each model includes controls for both union and nonprofit status. The top panel is for the entire sample. The middle and bottom panels have only non-managers and only managers, respectively. Table A1 repeats the ordered logits adding first a set of indicators for the establishment's primary

Table 1. Correlations between Job Design Attributes.

	Discretion		Skills	Interdependence
	Guidelines	Supervision		
Multitasking	0.8475	0.8505	0.8341	0.8485
Discretion				
Guidelines		0.8450	0.8234	0.8701
Supervision received			0.8274	0.8404
Skills				0.8176

Spearman's rank-order correlations between job design attributes. Because sample sizes are so large and significance levels are so high, those statistics are not shown in the tables. Overall sample size = 137,181; there are 15,349 firms, and 19,791 establishments.

Table 2. Unrestricted Relationships between Pairs of Job Design Attributes.

	Discretion		Skills	Interdependence
	Guidelines	Supervision		
(a) Full sample				
Multitasking	4.491	3.881	1.777	4.033
	(0.4759)	(0.4848)	(0.4218)	(0.4776)
Discretion				
Guidelines		3.395	1.470	3.756
		(0.4886)	(0.3916)	(0.5247)
Supervision			1.714	3.517
			(0.4308)	(0.4702)
Skills				2.952
				(0.3024)
(b) Non-managers only				
Multitasking	4.541	3.907	1.894	3.949
	(0.4538)	(0.4638)	(0.4120)	(0.4504)
Discretion				
Guidelines		3.901	1.566	3.684
		(0.4613)	(0.3887)	(0.5004)
Supervision			1.806	3.473
			(0.4201)	(0.4467)
Skills				3.039
				(0.2957)
(c) Managers only				
Multitasking	4.290	3.901	3.455	4.182
	(0.4283)	(0.4264)	(0.4147)	(0.4772)
Discretion				
Guidelines		4.568	2.774	4.016
		(0.4534)	(0.3255)	(0.5321)
Supervision			2.793	3.439
			(0.3605)	(0.4415)
Skills				3.028
				(0.3903)

Relationships between factors are coefficients from ordered logits; each cell represents a separate logit. Rows are dependent variables; columns are independent variables. Pseudo-R^2 are in parentheses. Additional controls included in each regression: union status and nonprofit status.

industry and then the job's primary occupation. Because of large sample sizes, all the coefficients have high levels of statistical significance, so standard errors are not included. A more informative statistic is the pseudo-R^2 (in parentheses below each coefficient): $1–(LL_{Full\ model}/LL_{Constant\ only})$,

where LL is the log-likelihood. The pseudo-R^2 shows the extent to which the variance in the dependent variable is "explained" by the model.

In all the models in the top panel of Table 2 for the full sample, the Pseudo-R^2 indicates a strong relationship between the factors. Close to half the variance in multitasking is explained by either of the discretion variables and by interdependence. Not surprisingly, there is also a strong positive relationship between the two measures of discretion. More than half the variance in guidelines is explained by interdependence. Overall, Table 2 presents strong evidence consistent with the prediction that job designs will tend to be "coherent" with respect to multitasking, discretion, and interdependence: these three characteristics are all positively associated with each other.

The relationships between skills and multitasking, skills and discretion, and skills and interdependence are also positive, but are not as strong. These suggest that, on balance, skills favor inter-task learning and continuous improvement rather than specialization. This is consistent with the evidence on skill-biased technological change and increasing returns to skill investments in recent decades. Rapid technological change reduces the incentive for firms to invest in ex ante optimization, and increases the opportunities for workers to make continuous improvements. That implies a trend toward multitasking and discretion. Our evidence suggests that these work even better if the worker has greater skills.

In addition to the results for the full sample at the top of Table 2, the results for the non-managerial and managerial samples are reported in the middle and bottom of the table. The first point of note is that the basic patterns are the same: strong positive correlations among all job design characteristics. Second, the correlations among skills and each of multitasking, guidelines, or supervision are much stronger within the managerial sample than within the non-managerial sample. This suggests that problem-solving skills are more valuable in managerial jobs.

That the evidence supports the theory for both the managerial and non-managerial samples, and the relationships are stronger when controlling for occupations, are particularly noteworthy in light of previous empirical evidence. The examples studied most often come from manufacturing, and are closely tied into the discussion in recent years of the impact of human resource practices on productivity and profitability (Huselid, 1995; MacDuffie, 1995; Ichniowski et al., 1997; Cappelli & Neumark, 2001). The disproportionate focus on manufacturing is understandable given the intellectual heritage and framework established by Taylor (1923), and the ease of measuring productivity in manufacturing. But the theory does not

require a manufacturing setting, as the more recent research on service environments demonstrates (Batt, 2002; Batt & Moynihan, 2002).[8] Yet despite the gains that have been made at the case study level, to date there has been no systematic data available to test these predictions economy-wide. Table 2 provides the first such evidence.

4.2. Multivariate Relationships between Job Characteristics

The results in Tables 2 and A1 provide evidence that pairs of job design attributes – including skills – are complementary. A stronger test focuses on the extent to which they cluster together as a group so that job designs are "coherent" at the job level – all dimensions high, all medium, or all low – which we test in Table 3. At the top of Table 3 are the distributions of each dimension relative to the median in the entire sample.[9] Because we expect that occupations segregate jobs into groups that are already similar on each job design dimension, we want to focus on the extent to which a job is low, medium, or high relative to the occupational norm. Consequently, in the second panel of Table 3 we center the values for each job around the median for each 3-digit occupation. Comparing the patterns in the top two panels of Table 3, the much higher concentration at the median in the second panel shows that occupations group together jobs that are similar along each design dimension.

To construct a multidimensional measure to test whether job design dimensions group together as all high, all low, or all medium along all four dimensions, we first use the rankings in the middle panel of Table 3 to assign a value of 1 (below the occupational median = L), 2 (at the occupational median = M), or 3 (above the occupational median = H) to each job for each dimension. We then sum these values for each job to create an index that ranges from 4 (LLLL) to 12 (HHHH) for each job. There are 81 possible combinations of the four characteristics, and 9 possible sums. The bottom panel of Table 3 shows the percentage of jobs with all low values, all high, all medium, as well as all other possible sums. The value of 8 is broken into two groups: jobs that have all medium (MMMM) for all four dimensions, and those that have an index value of 8 via some other combination of values (e.g., LHMM, MLMH, HMML, etc.). The first column contains the actual distribution of the index values in the sample, with the standard error of each percentage in parentheses under the mean. The second column has the probability that that index should occur if the

Table 3. Distribution of Leveling Factors.

	L (<Median)	M (Median)	H (>Median)
Distribution relative to median value in the economy			
Skills	0.362	0.199	0.439
Guidelines	0.333	0.361	0.306
Multitasking	0.193	0.351	0.456
Interdependence	0.309	0.345	0.346
Distribution relative to median value within 3-digit occupation			
Skills	0.251	0.540	0.209
Guidelines	0.190	0.610	0.200
Multitasking	0.194	0.603	0.203
Interdependence	0.185	0.619	0.196

Index Relative to Median	Fraction of All Jobs (SE)	Pr(Characteristics Randomly Assigned from Empirical Distribution)	Actual/Predicted
Index (Σ) of skills, guidelines, multitasking, and interdependence (using distribution relative to median value within 3-digit occupation)			
4 (= LLLL)	0.0541 (0.0006)	0.0017	31.6
5	0.0697 (0.0007)	0.0202	3.4
6	0.1109 (0.0009)	0.0957	1.2
7	0.1488 (0.0010)	0.2320	0.6
8 (= MMMM)	0.2502 (0.0012)	0.1230	2.0
All other values of index = 8 except MMMM	0.0151 (0.0003)	0.1856	0.1
9	0.1268 (0.0009)	0.2278	0.6
10	0.0796 (0.0007)	0.0929	0.9
11	0.0823 (0.0007)	0.0196	4.2
12 (= HHHH)	0.0626 (0.0007)	0.0017	37.6

values in the middle panel of the table were randomly distributed across all jobs. The third column has the ratio of the actual to predicted values.

The strong test of the extent to which firms choose between classical and modern job designs across jobs is provided by comparing the percentage of

jobs with all low or all high values to the expected percentage if job characteristics were randomly assigned based on their univariate frequency distributions from the middle panel of Table 3. For example, the expected percentage of workers with all low values equals the product of the percentages of jobs below the median for each characteristic: $(0.251) \times (0.190) \times (0.194) \times (0.185) = 0.0017$ (third column). The corresponding expected percent having all high values is, coincidentally, also 0.0017. The actual occurrence of both job types (LLLL and HHHH) is more than 30 times more likely than one would expect purely by chance. The actual occurrence of MMMM jobs is not as dramatic relative to the random case, but is still quite divergent – twice as likely. Moreover, jobs that are "almost all high" (index value of 11, which means three H and one M) or "almost all low" (index value of 5, which means three L and one M) occur three to four times as often as is expected by chance. Thus the patterns in the bottom panel of Table 3 provide strong evidence of coherence in job design at the individual job level.

It is worth asking whether the percentages of jobs falling into the high and low groups are what we would expect to see, given all that has been written about the trends toward modern job design. The fraction of all jobs that have an index value of either 11 ("almost all high") or 12 ("all high"), which we view as a reasonable proxy for modern jobs, is 14.5 percent. Similarly, the fraction of all jobs that have an index value of either 4 ("all low") or 5 ("almost all low"), which we view as a reasonable proxy for classical jobs, is 10.4 percent. Given that these are the first nationally representative data available with these types of job design measures, it is difficult to determine whether such a figure should be viewed as high, low, or "just right."

On one hand, 14.5 percent modern jobs suggests that the trend toward modern job design has not been very pervasive, though it is up to the reader to decide if almost one out of six jobs is a relatively large fraction, given that we have only cross-sectional data. On the other hand, the percentage of all jobs that are modern (combining index values 11 and 12) is proportionately almost one and a half times larger than the percentage of all jobs that are classical (combining index values 4 and 5). Viewed this way, it would appear that modern jobs are more pervasive than classical jobs, supporting the claim that there may have been a shift toward modern jobs in recent years.

Perhaps more interesting is the percentage of jobs with design attributes that cluster at the middle (index value of 8 and MMMM). At 25 percent, these account for one-quarter of all jobs, which is twice as prevalent as one would expect if job design attributes were chosen at random. While the literature and business press has focused predominantly on the two

extremes – modern versus classical jobs – there is some evidence that firms face difficulties when attempting to implement modern job design. For example, teams are one example of modern job design that combines cross-functional responsibilities (multitasking), training (higher skills), and decentralized decision making (discretion, low supervision). The literature on teams is replete with evidence that they are difficult to set up, administer, and maintain (Mohrman, Cohen & Mohrman, 1995; Osterman, 2000; Gibson & Cohen, 2003). Thus organizations may be in a continual state of flux with respect to teams, sometimes expanding their use and sometimes contracting as they struggle to implement them effectively (Levenson, 2007).

If efforts at implementing teams and modern jobs often fall short of their intended goal, the end result could be a number of jobs that are more modern than classical, but "not quite modern enough" to fit the ideal as characterized by the "all high" jobs in Table 3. This is consistent with the evidence documented by Ichniowski and Shaw (1995) that firms tend to adopt clusters of HR practices that are consistent with modern job design, but that there are costs of adoption due to changing over from classical to modern job design. This leads new establishments to be more likely to adopt the most wide-reaching sets of HR practices and modern job design, while older establishments (that start off with more classical job designs) are more likely to adopt some, but not all, such modern job design practices. While we are unable to match our job design data to comparable measures of HR practices, the evidence in Table 3 of a disproportionately large number of "medium" (MMMM) jobs is consistent with a cost of adoption story.

The rest of the analysis in the chapter uses the classification of jobs into LLLL, MMMM, and HHHH categories defined in Table 3. As such, it is worth emphasizing what those classifications represent. Note that in each case the classification is *relative to the 3-digit occupation median*. Standard occupational classifications inherently represent job characteristic clustering: professional and technical occupations tend to be high on knowledge, discretion, and complexity, while manual labor jobs and administrative support occupations tend to be lower on knowledge, discretion, and complexity.

From a production standpoint, firms do not have a huge amount of latitude to substitute across broad occupation categories when designing jobs: lawyers and scientists typically cannot be substituted en masse for secretaries, laborers, and truck drivers without introducing massive distortions in the marginal cost of production (wage costs) and/or efficiency of production. The job design model focuses on the decisions firms make within occupation categories: the extent to which knowledge, discretion, and

complexity vary *within* technical/professional jobs, or *within* manual labor jobs, or *within* administrative support jobs. The interesting question is what leads some scientists' jobs to have greater complexity, knowledge and/or discretion relative to other scientists' jobs; and the same thing for truck drivers relative to other truck drivers, secretaries relative to other secretaries, etc. Our analysis focuses on those job design decisions, not the decision of how many jobs of different occupational types the firm needs.

4.3. Effects of Establishment Characteristics on Job Characteristics

We have argued that no single job design strategy is optimal for all types of establishments, but that characteristics of the environment, such as product complexity, stability, and predictability will affect the choice of job design. We start by examining whether unionization, establishment size, and nonprofit status affect job design, modeling the probability that a job is "all modern" or "all classical" using logit regressions. Table 4 shows the results of this analysis.[10] The second and fourth columns include a full set of industry indicators.

Unionized jobs are much less likely to be "all classical" yet also less likely to be "all modern." The former is consistent with unions' traditional negative views of classical job design. The latter is consistent with the

Table 4. Determinants of Modern (HHHH) or Classical (LLLL) Job Design.

	Pr(LLLL)	Pr(LLLL)	Pr(HHHH)	Pr(HHHH)
Nonprofit	−0.1115	−0.2911**	−0.2193**	−0.2303**
Union	−0.8562*	−0.7078*	−0.1755*	−0.1801*
Employment/1,000	−0.0226*	0.0054	0.0820*	0.0387*
(Employment/1,000)^2	−0.0001**	−0.0001	−0.0011*	−0.0003*
Industry controls	No	Yes	No	Yes
Pseudo-R^2	0.0128	0.0679	0.0109	0.0817
N	42,750	41,586	42,750	41,870

Coefficients from logits. Sample = jobs in multi-establishment firms. Controls are included for percent of jobs in 14 job design clusters as described in Table 7a.
*p-value < 0.01.
**p-value < 0.05.

conventional wisdom that unions resist change, and to wider differences in compensation among nonmembers. Modern job design has potential benefits to employees in upgraded skills and potentially higher wages. But making that change can threaten the probability that existing union workers will keep their jobs, and might widen the dispersion in earnings among members. Nonprofits similarly reduce the probability that a job is either ''all modern'' or ''all classical.''

Larger establishments are more likely to choose modern job design and less likely to choose classic job design. This is consistent with the model, which argues that multitask output can exceed specialized output when coordination costs are large. In larger establishments there are often more hierarchical levels, making information transfer slower and more difficult, resulting in higher coordination costs. Finally, it is important to note that although these establishment characteristics alone do not explain a large fraction of the variance in the probability a job is modern or classical, the industry indicators add substantial explanatory power to the model. This suggests that other characteristics of the industry, such as product complexity and stability, do strongly affect an establishment's choice of job design.

One critique of our findings might be that they are driven not by inter-task learning, but instead by firms designing jobs to generate intrinsic motivation as in the social psychology literature. The fact that job design patterns vary systematically across different industries suggests that product or industry characteristics matter, which can be taken as evidence in favor of the inter-task learning explanation. However, we do not have sufficient data to rule out the possibility that the returns to generating intrinsic motivation through job design do not vary by industry. In a model such as Lindbeck and Snower's (2000), differences in worker preferences for modern jobs and ability to multitask are two of the factors that lead to observed differences in the adoption of modern job design. If the supply of workers to particular industries is determined in part by sorting on noneconomic preferences for working in those industries (e.g., an intrinsic preference for social work or education versus manual labor), and if workers' multitasking ability is related to those preferences, then firms in those different industries may face differential returns to using job design to tap into intrinsic motivation. Of course, it is most likely that both mechanisms play a role: there may be differential returns to generating intrinsic motivation at the same time that product/industry characteristics are an important determinant of the returns to adapting modern job design.

4.4. Technology and Job Design

While we do not have direct measures of industry characteristics such as product complexity and stability available in the NCS data, we were able to match the NCS job design characteristics at the industry level to measures of aggregate computer use and R&D spending to investigate the interaction of technology and job design.

Table 5 focuses on the relationship between job design choices (modern versus classical) and computer usage. The computer usage data comes from the September 2001 Internet and computer use supplement to the current population survey, and are matched at the 2-digit industry level using the CPS microdata. This enabled matching for 51 distinct industry groupings. Two sets of correlations with computer usage are presented: the percentage of jobs in an industry that are modern, and the percentage that are classical. In both cases the correlations using both percentages and ranks are presented at the bottom of the table. Computer usage and the percentage of jobs in the industry that are modern are fairly strongly correlated (0.50), indicating that computerization and the design of jobs to deal with complexity, interdependence, and autonomy are closely related, consistent with computers being a complement to skill, at least for some jobs. Computer usage is also positively correlated with the percentage of jobs in an industry that are classical (0.30), consistent with computers being used to increase monitoring, decrease autonomy, and lower the skill requirements for other jobs. These patterns are consistent with industries using computers to simultaneously upskill some jobs while downskilling other jobs (Goldin & Katz, 1998; Autor et al., 2002).

Table 6 shows the relationship between R&D spending and job design. The R&D data come from NSF, Division of Science Resources Statistics, Research and Development in Industry: 1999, NSF 02-312. R&D spending per capita was calculated using the aggregate employment for each industry from that same source. Accurate R&D numbers are not available at the same level of disaggregation as the computer usage data, hence, there are only 17 industries available for this analysis. Despite the small sample size, the correlation of per capita R&D spending with the percentage of jobs that are modern in an industry is very high (0.76) and is statistically significant. The correlation of per capita R&D spending with the percentage of jobs that are classical, in contrast, is both much smaller (0.20) and not statistically significant. Given the small sample size, it may be the case that a larger sample would produce different correlation patterns. Thus the results in Table 6 should be taken as preliminary evidence that R&D spending is

Table 5. Computer Usage and Industry Patterns of Job Design.

Industry	% Using Computers at Work	Rank	% Jobs Modern	Rank	% Jobs Classical	Rank
Brokers	0.912	1	0.120	1	0.056	33
Mfg. – Prof. Equipment	0.728	10	0.117	2	0.058	29
Mfg. – Chemicals	0.720	12	0.117	3	0.036	42
Service – Professional	0.837	6	0.115	4	0.096	4
Mfg. – Transport	0.509	28	0.112	5	0.037	41
Mfg. – Machine	0.584	21	0.096	6	0.050	37
Mfg. – Paper	0.492	30	0.087	7	0.031	46
Service – Legal	0.882	2	0.086	8	0.097	3
Mfg. – Stones	0.396	42	0.084	9	0.052	35
Mining	0.448	37	0.079	10	0.065	21
Insurance	0.867	3	0.077	11	0.064	23
Mfg. – Electric	0.621	19	0.073	12	0.078	13
W. Durables	0.641	17	0.072	13	0.085	8
Mfg. – Petroleum	0.844	5	0.067	14	0.094	6
Utility	0.608	20	0.066	15	0.034	45
Service – Nonprofessional	0.677	14	0.065	16	0.065	22
W. Nondurables	0.526	27	0.064	18	0.078	14
Public Administration	0.746	9	0.064	17	0.051	36
Mfg. – Printing	0.630	18	0.063	20	0.091	7
Real Estate	0.652	16	0.063	19	0.081	10
Service – Entertainment	0.472	33	0.062	21	0.048	38
Banking	0.853	4	0.060	22	0.075	15
Retail – Catalog	0.537	24	0.059	24	0.116	1
Mfg. – Rubber	0.477	32	0.059	25	0.063	25
Service – Social	0.446	38	0.059	26	0.062	26
Communications	0.812	7	0.059	23	0.060	27
Mfg. – Food	0.334	45	0.057	27	0.069	19
Mfg. – Metal	0.471	34	0.055	28	0.057	30
Transport	0.405	40	0.055	29	0.036	43
Service – Education	0.701	13	0.054	30	0.060	28
Mfg. – Toys, etc.	0.467	35	0.050	31	0.071	17
Service – Hotel	0.416	39	0.050	32	0.019	49
Retail – Gas	0.451	36	0.046	33	0.017	51
Mfg. – Lumber	0.298	48	0.041	34	0.057	31
Service – Hospital	0.722	11	0.037	35	0.036	44
Construction	0.256	50	0.035	37	0.072	16
Retail – Grocery	0.346	44	0.035	36	0.019	50
Service – Business	0.657	15	0.032	38	0.095	5
Retail – Vehicle	0.559	22	0.032	39	0.048	39
Service – Medical	0.549	23	0.032	40	0.04	40
Retail – Eating	0.240	51	0.029	41	0.027	48
Mfg. – Leather	0.531	26	0.028	42	0.067	20

Table 5. (*Continued*)

Industry	% Using Computers at Work	Rank	% Jobs Modern	Rank	% Jobs Classical	Rank
Mfg. – Apparel	0.282	49	0.027	44	0.064	24
Retail – Building	0.491	31	0.027	43	0.055	34
Retail – Other	0.493	29	0.026	46	0.109	2
Retail – Hobby	0.535	25	0.026	45	0.071	18
Mfg. – Textile	0.300	47	0.024	47	0.057	32
Service – Repair	0.378	43	0.023	48	0.079	11
Service – Personal	0.316	46	0.022	49	0.083	9
Retail – Apparel	0.403	41	0.014	50	0.029	47
Retail – Technology	0.752	8	0.006	51	0.079	12
Correlation between computer use and % jobs modern and classical:			0.50***	0.51***	0.30***	0.29***

highly complementary with modern job design, and much less complementary (if not unrelated) to classical job design. This is consistent with R&D spending being focused on innovations that increase product complexity and that require processes that are optimized when workers have greater autonomy and skills. Moreover, organizations that invest more in R&D tend to have greater opportunities for continuous improvement. Their industries are more likely to involve rapid technological change and unpredictability, so that ex ante optimization is less effective.

The combined results in Tables 5 and 6 provide good evidence that job design decisions are related to a firm's or industry's product characteristics and technology. There is one additional point about the patterns in Tables 5 and 6 that is worth noting. At first glance, the reader might find the prevalence of modern versus classical jobs in certain industries to be anomalous. For example, professional services has the highest rate of classical jobs in Table 6 and the fourth highest rate of classical jobs in Table 5. Professional services is commonly thought of as an industry in which discretion and customized work are widespread. Thus it might appear counterintuitive to find that this industry has a disproportionately large fraction of classical jobs. However, there are two reasons why this finding is not necessarily wrong and may in fact be reasonable.

First, note that no industry has more than about 10 percent purely classical jobs (all low, LLLL) in either table. Given the large number of tasks and jobs involved in any industry, even if the typical job in an industry might be modern, there is nothing preventing a minority of other jobs from

Table 6. R&D Spending and Industry Patterns of Job Design.

Industry	R&D ($Millions)	Domestic Employment (Thousands)	R&D per Thousand Employees	Rank	% Jobs Modern	Rank	% Jobs Classical	Rank
Mfg. – Chemicals	20,372	1,023	19.91	2	0.117	1	0.036	16
Service-Professional	23,640	761	31.06	1	0.115	2	0.096	1
Mfg. – Transport	34,059	2,159	15.78	4	0.112	3	0.037	14
Mfg.-Machine, Prof. Eq.	44,076	2,230	19.77	3	0.102	4	0.052	12
Durables, Nondurables	19,960	1,339	14.91	5	0.068	5	0.081	3
Mfg. – Petrol	615	116	5.30	9	0.067	6	0.094	2
Utility	142	410	0.35	17	0.066	7	0.033	17
Communications	15,421	1,665	9.26	8	0.059	8	0.060	8
Mfg. – Rubber	1,845	562	3.28	10	0.059	9	0.063	7
Mfg. – Food	1,159	1,043	1.11	13	0.057	10	0.069	6
Transport	466	756	0.62	16	0.055	11	0.037	15
Mfg. – Metal	2,174	1,120	1.94	12	0.055	12	0.057	11
Mfg. – Toys, etc.	4,226	351	12.04	7	0.050	13	0.071	5
Mfg. – Lumber	70	71	0.99	14	0.041	14	0.057	10
Construction	699	270	2.59	11	0.034	15	0.072	4
Svc. – Med., Hospital	660	51	12.94	6	0.034	16	0.039	13
Mfg.-Textile, Apparel, Leather	337	362	0.93	15	0.026	17	0.059	9
Correlation between per capita R&D spending and % jobs modern and classical:					0.76***		0.20	

being classical – particularly given the tendency of firms to segregate modern and classical jobs by establishment, as is shown in the next section below. An industry such as professional services that often produces highly customized products may be able to do so by employing one set of people who focus on the higher order knowledge work (and work in more modern jobs), while simultaneously employing a second set of people who focus solely on routine work (and work in more classical jobs). The former could include the client-facing work and most difficult problem-solving tasks that involve generating new knowledge and work routines for the clients, while the latter could include more routine tasks such as order processing, data entry, filing regulatory compliance forms, managing billing processes, etc.

Second, recall that we defined a job as high versus medium versus low on each job design dimension *relative to its occupational norm*. Thus an industry that typically employs college graduates in knowledge work jobs will only show up as having a large fraction of modern jobs if the amount of discretion, skills, etc. is high *relative to comparable jobs in other industries that also require college graduate-level skills to perform knowledge work*. Thus the patterns in Tables 5 and 6 are driven by relative differences in concentrations of modern versus classical jobs within, not between, occupations. A ranking of industries on the basis of job design characteristics that did not account for occupational norms would naturally give greater weight to industries employing larger number of more highly educated workers, who tend to have jobs with greater discretion in general. For our purposes, we find the rankings in Tables 5 and 6 to be more useful and informative.

4.5. Similarity of Job Designs within Firms and Establishments

We now analyze the prediction that job designs will tend to be similar within firms, and even more so within establishments. The relevant comparison for a job is not to all other jobs in the economy, but to other jobs in the same establishment or firm. We reestimate the logits of the previous section, including as regressors the percentages of other jobs in the establishment or firm that fall into each of the 81 unique combinations of the four job characteristics. For ease of interpretation, Table 7A reports the results when all jobs with common combinations are grouped together. For example, the "3L, 1M" group includes four subgroups: LLLM, LLML, LMLL, and MLLL.[11] We predict that the probability that any one job is "all modern" is positively related to how many other jobs in the establishment and/or the

Table 7A. Effect of Distribution of Other Jobs' Characteristics on Probability of Modern (HHHH) or Classical (LLLL) Job Design.

Skill Set	Pr(LLLL)		Pr(LLLL)		Pr(HHHH)		Pr(HHHH)	
	Establishment	Firm	Establishment	Firm	Establishment	Firm	Establishment	Firm
LLLL	**3.281***	**3.078***	**2.376***	**2.101***	**−0.547****	0.588	**−0.931***	0.191
3L, 1M	**1.262***	**1.045***	**1.199***	**0.848***	−0.395	−0.410	**−0.603****	**−0.615*****
2L, 2M	**1.176***	0.015	**1.158***	−0.120	**−0.572***	−0.408	**−0.585***	−0.417
1L, 3M	**0.520***	0.175	**0.539***	0.076	**−0.758***	**−0.800***	**−0.732***	**−0.670***
3L, 1H	2.104	5.729	1.477	3.907	−0.996	−2.412	−1.590	−4.747
2H, 1M, 1L	0.870	1.696	0.494	1.016	0.495	0.611	0.495	0.779
2H, 2L	−0.888	0.042	−4.151	−3.046	−11.67	7.032	−10.70	6.922
1L, 2M, 1H	−0.472	1.047	−0.390	1.038	0.073	−0.148	−0.243	−0.730
2L, 1M, 1H	−1.639	0.145	−1.289	0.430	−0.171	0.977	−1.252	−0.120
3H, 1L	0.426	−0.571	0.726	−0.115	−2.658	−2.720	−2.251	−1.369
1H, 3M	**−0.986***	−0.059	**−0.948***	0.234	**0.427***	−0.361	**0.431***	−0.109
2H, 2M	**−0.841***	−0.372	**−1.101***	**−0.948***	**1.019***	0.084	**0.887***	−0.202
3H, 1M	**−0.624****	0.495	**−0.946***	0.087	**1.354***	0.200	**1.275***	0.303
HHHH	**−0.816***	**0.970***	**−1.194***	0.338	**3.517***	**1.937***	**2.799***	**1.159***
Industry controls	No		Yes		No		Yes	
Pseudo R^2	0.0926		0.1143		0.1133		0.1297	
N	41,421		40,285		41,421		40,570	

Coefficients from logits. Controls included for nonprofit status, unionization, establishment size and its square. Sample = jobs in multi-establishment firms.

*p-value < 0.01.

**p-value < 0.05.

***p-value < 0.10.

firm are "all modern." For the firm variables, the percentages are calculated using jobs at other establishments in the same firm, excluding jobs at the same establishment. Thus firms with only one establishment are excluded from the analysis in Table 7A. The first set of columns predicts the probability of a classical (LLLL) job, both with and without 3-digit industry controls. The second set of columns predicts the probability of a modern (HHHH) job.

The results in Table 7A are consistent with the predictions. The probability of a classical job is correlated positively with the percentage of other jobs in the establishment that are classical (first row), and negatively with the percentage of other jobs in the establishment that are modern (last row). Similarly, the probability of a modern job is correlated positively with the percentage of other jobs in the establishment that are modern, and negatively with the percentage of other jobs in the establishment that are classical. There are similar positive, but smaller, correlations between Pr(LLLL) and many of the jobs that are "almost all" classical (3L1M) and "mostly classical" (2L2M; 1L3M). The opposite is true for Pr(HHHH) and jobs that are almost (3H1M) or mostly (2H2M; 1H3M) modern. Jobs that mix both high and low characteristics (3L1H; 2L2H; 1L2M1H; etc.) are much less likely to be positively correlated with either Pr(LLLL) or Pr(HHHH): none of those coefficients have p-values < 0.05. Thus, firms tend to choose pure job design approaches, opting for many jobs to be either high on all dimensions, or low on all dimensions.

To a lesser degree, firms make the same choice across establishments, as predicted. This provides evidence that respondent bias is not the explanation for correlations between job designs with those of other jobs in the establishment. Although we are concerned that a single human resource representative describing all sampled jobs in the establishment may scale up or down all responses, jobs across establishments within a single firm are described by separate individuals. If job design were not clustered within an establishment but merely appeared to be so due to respondent bias, we would not expect to find peer effects for other workers within the firm but outside the establishment – such effects confirms that respondent bias is not driving the results.[12] Patterns in job design within industries and occupations, described below, are further evidence that our findings are not driven by respondent bias.

Two additional patterns are worth noting in Table 7A. First, having many modern jobs in the same establishment reduces the probability that a job will be classical. At the same time, having a high percentage of modern jobs in the *other* establishments in the firm increases the probability that

a job will be classical in the present establishment, too. This suggests that firms isolate similar jobs in the same establishment and also push job design toward the extremes, away from the middle. This pattern disappears when controlling for industry differences across establishments. Thus, such clusters of establishments are concentrated in some industries and not others, and this pattern likely is related to differences in product, technology and/or organizational change.[13]

Second, some within-establishment correlations get *stronger* when controlling for industry. Specifically when predicting Pr(LLLL), coefficients on the fraction of jobs that are HHHH and (3H1M) get more negative; and when predicting Pr(HHHH) coefficients on the fraction of jobs that are LLLL and (3L1M) get more negative. This means that the tendency for a firm to segregate modern and classical jobs across its establishments is consistent across industries, though more prevalent in some industries.

Table 7B presents the results from predicting Pr(MMMM), using the same set of regressors as Table 7A. As expected, the probability that a job will be MMMM is strongly correlated with the presence of similar "all medium" jobs in both the establishment and in the firm, with stronger within-establishment than within-firm correlations. Table 7B shows the same within-firm, across-establishment segregation of dissimilar jobs. In the case of "medium" jobs in Table 7B, the segregation occurs for jobs that are only slightly different. For example, the greater the fraction of (1H3M) jobs in the rest of the firm, the lower the probability of a MMMM job in the same establishment.

4.6. Within versus Outside 2-Digit Occupation Correlations

To this point, we have not distinguished between occupations except to control for nationwide differences in the median value for each leveling factor by occupation. An interesting question is the extent to which job design patterns within an establishment are driven by clustering of jobs in similar occupations, where occupations are defined by Census 2-digit classifications. We would expect some within-2-digit-occupation clustering, given task interdependencies and the consequent complementarity of such skills in production; for example, grouping modern chemical engineering with modern electrical engineering jobs. Less obvious is the prediction of between-2-digit-occupation clustering; for example, grouping modern engineering with modern administrative support jobs. It is reasonable to expect such clustering if the task interdependencies in production are

Table 7B. Effect of Distribution of Other Jobs' Characteristics on Probability of MMMM Job Design.

Skill Set	Pr(MMMM)		Pr(MMMM)	
	Establishment	Firm	Establishment	Firm
LLLL	0.059	**−0.453****	0.047	**−0.617***
3L, 1H	**−5.404***	−0.203	**−5.452***	0.250
2H, 1M, 1L	−0.386	−0.361	−0.522	−0.678
2H, 2L	−0.321	1.880	−0.396	1.099
3L, 1M	**0.338****	**−0.526***	0.280	**−0.654***
2L, 2M	**0.331***	**−0.541***	0.193	**−0.793***
1L, 3M	**0.422***	−0.188	**0.279****	**−0.490***
MMMM	**1.508***	**1.237***	**1.185***	**0.669***
1H, 3M	**0.465***	**−0.501***	**0.278****	**−0.959***
2H, 2M	**0.459***	**−0.467***	**0.454***	**−0.535***
3H, 1M	**0.427***	**−0.777***	**0.347***	**−1.046***
1L, 2M, 1H	−0.068	**−0.717****	−0.007	**−0.746****
2L, 1M, 1H	−0.850	−1.261	−0.872	−1.349
3H, 1L	0.283	1.307	−0.050	0.930
Industry Controls	No		Yes	
Pseudo R^2	.0406		.0459	
N	41,421		41,298	

Coefficients from logits. Controls included for nonprofit status, unionization, establishment size and its square. Sample = jobs in multi-establishment firms.
*p-value < 0.01.
**p-value < 0.05.

relatively "global" across the entire production process. For the most peripheral tasks, however, we would expect interdependencies to diminish to the point where there are fewer gains from clustering job design attributes; such tasks likely would include non-"core" processes such as janitorial work and food service. One characteristic of truly peripheral tasks is that they should be greater candidates for outsourcing (Abraham & Taylor, 1996).

Table 8 shows the proportion of jobs outside of one's own occupation that have the same job design (a) for the economy absent one's own firm, (b) for the firm absent one's own establishment, and (c) for the other jobs in the establishment. For the sample of single-establishment firms, only the first and third categories are relevant. The clustering of modern and classical jobs is greater at the establishment level than at the firm level and in the economy overall: both modern and classical jobs are approximately twice as

Table 8. Clustering of Job Design Outside Own 2-Digit Occupation.

Job	Proportion of jobs outside own occupation with same job design				
	Multi-Establishment Firms			Single-Establishment Firms	
	Like jobs in economy absent own firm	Like jobs in firm absent own establishment	Like jobs in establishment absent own firm	Like jobs in economy absent own firm	Like jobs in establishment absent own firm
LLLL	0.0525	0.0684	0.0949	0.0526	0.0967
MMMM	0.2482	0.2536	0.2513	0.2481	0.2460
HHHH	0.0618	0.1292	0.1604	0.0620	0.1132

likely to be observed within an establishment as in the economy at large. This confirms our findings in Tables 7A and 7B and suggests that occupational clustering intrinsic to the production process does not entirely drive the job design clustering results. For classical (LLLL) jobs, the establishment-level clustering is the same at single- versus multi-establishment firms. For modern (HHHH) jobs, the establishment-level clustering is much stronger in multi-establishment firms. Thus larger (multi-establishment) firms are much more likely to cluster dissimilar modern jobs together. The degree of clustering of all "medium" jobs, in contrast, is no greater within-firm or within-establishment than in the economy overall.

In Table 9, we perform a more rigorous test of the relative importance of within- and across-occupation clustering of job design, by reestimating the models in Table 7A, separating each within-establishment job design variable into two components: similarly designed jobs within the same occupation and similarly designed jobs in all other occupations. The results show there is both within- and across-2-digit-occupation clustering of job design types at the establishment level. For modern jobs, the coefficients on the percentage of other jobs in the establishment that are modern both within the same 2-digit occupation and in other 2-digit occupations are positive and significant at the $p<.01$ level (bottom row, fourth and fifth columns). The pattern is the same for classical jobs (top row, first two columns). Moreover, in both cases the within-2-digit-occupation correlation is stronger than the across-2-digit-occupation correlation, indicating that within-occupation clustering is more likely than across occupation clustering, as expected. More important is the fact that across-occupation clustering drives at least part of the results in Table 7A: firms tend to group together jobs that are all modern and all classical, even dissimilar jobs.

Table 9. Effect of Distribution of Other Jobs' Characteristics on Probability of Modern (HHHH) or Classical (LLLL) Job Design: Comparing Jobs Within and Outside Own 2-Digit Occupation.

Peers in Skill Set	Pr(LLLL)			Pr(HHHH)		
	Jobs in the establishment		Jobs in other establishments in firm	Jobs in the establishment		Jobs in other establishments in firm
	Within own 2-digit occupation	Outside own 2-digit occupation		Within own 2-digit occupation	Outside own 2-digit occupation	
LLLL	**1.8851**	**0.6491**	**2.1121**	**−0.724***	**−0.856***	0.263
3L, 1M	**0.5401**	**0.8971**	**0.9661**	**−1.375***	−0.352	−0.463
2L, 2M	**0.5051**	**0.7031**	−0.010	**−0.953***	**−0.364****	−0.193
1L, 3M	−0.207	**0.3875**	0.169	**−0.463***	**−0.683***	**−0.547****
3L, 1H	0.706	3.504	3.274	1.012	−0.771	−4.905
2H, 1M, 1L		0.514	1.326	−0.324	−0.363	1.062
2H, 2L	−1.004	−2.768	−3.699		−10.14	7.326
1L, 2M, 1H	0.078	0.065	1.297	**−0.944****	−0.235	−0.912
2L, 1M, 1H	−1.091	−1.427	0.855	−0.307	−1.088	0.333
3H, 1L	−2.098	1.199	−0.322	1.057	−2.071	−0.619
1H, 3M	**−0.935***	**−0.569***	0.278	−0.018	0.237	0.051
2H, 2M	**−0.560***	**−0.866***	**−0.955***	**0.4241**	**0.4831**	−0.110
3H, 1M	**−0.852***	**−0.735***	0.188	**0.2651**	**0.7301**	0.485
HHHH	**−0.7511**	**−0.9661**	0.379	**1.9481**	**0.8711**	**1.2571**
Industry controls		Yes			Yes	
R^2		0.13			0.15	
N		39,519			39,806	

Results from logits. Sample = jobs in multi-establishment firms.
*p-value < 0.01.
**p-value < 0.05.

To better understand these dynamics, Table 10 presents the analog of Table 8 for modern and classical jobs in multi-establishment firms for each of the 2-digit-occupation classifications. This enables an identification of which types of jobs drive the across-occupation clustering results in Table 9. For example, using the overall mean in the first row of column three as the comparison, the occupations for which modern jobs are more likely to be clustered with modern jobs in dissimilar occupations at the establishment

Table 10. Clustering of HHHH and LLLL Job Design Outside Own 2-Digit Occupation.

	Proportion of other jobs with same job characteristics mix					
	LLLL			HHHH		
	All jobs in economy, not in firm	All jobs in firm, not in establishment	All other jobs in establishment	All jobs in economy, not in firm	All jobs in firm, not in establishment	All other jobs in establishment
All workers	0.053	0.068	0.095	0.062	0.129	0.160
Public administration	0.054	0.047	0.283	0.062	0.081	0.105
Executives	0.052	0.053	0.085	0.058	0.132	0.173
Management related	0.051	0.069	0.096	0.062	0.209	0.258
Engineers	0.053	0.061	0.093	0.061	0.172	0.229
Math/computer science	0.053	0.062	0.104	0.062	0.363	0.408
Natural science	0.054	0.062	0.101	0.062	0.157	0.204
Health diagnostic	0.054	0.064	0.065	0.063	0.079	0.100
Health treatment	0.055	0.062	0.094	0.065	0.071	0.053
University professor	0.054	0.074	0.068	0.062	0.082	0.103
Teachers	0.054	0.036	0.181	0.065	0.033	0.118
Lawyer/judge	0.054	0.050	0.095	0.062	0.066	0.154
Other professional	0.053	0.084	0.101	0.063	0.124	0.182
Health technology	0.054	0.045	0.071	0.063	0.092	0.090
Engineering technology	0.054	0.041	0.082	0.063	0.189	0.241
Other technology	0.053	0.077	0.077	0.062	0.185	0.171
Sales manager				0.062	0.045	0.030
Finance/business ales	0.054	0.093	0.129	0.062	0.032	0.056
Service sales	0.054	0.023	0.013	0.062	0.346	0.347
Retail sales	0.056	0.135	0.167	0.065	0.082	0.117
Other sales	0.054	0.104	0.090	0.063	0.029	0.092
Admin. supervisor	0.055	0.004	0.000	0.063	0.138	0.171

Computer operator	0.054	0.058	0.063			
Secretary	0.055	0.156	0.197	0.062	0.101	0.152
Records	0.054	0.111	0.120	0.063	0.113	0.110
Mail distribution				0.063	0.024	0.086
Other admin.	0.050	0.074	0.099	0.064	0.135	0.168
Protective services	0.055	0.076	0.069	0.063	0.107	0.126
Food services	0.057	0.018	0.000	0.064	0.054	0.070
Health services	0.056	0.000	0.000	0.063	0.145	0.144
Building services	0.056	0.007	0.042	0.061	0.084	0.095
Personal services	0.055	0.026	0.025	0.061	0.082	0.059
Mechanic	0.054	0.059	0.076	0.065	0.191	0.167
Construction	0.054	0.034	0.023	0.064	0.120	0.197
Other precision	0.054	0.073	0.088	0.064	0.117	0.205
Machine operator	0.053	0.031	0.078	0.062	0.117	0.181
Assembler	0.055	0.047	0.116	0.062	0.091	0.127
Vehicle operator	0.054	0.116	0.053	0.063	0.135	0.115
Other transportation	0.054	0.047	0.169	0.063	0.083	0.182
Construction laborer				0.062	0.089	0.043
Handlers				0.061	0.069	0.075
Other laborer				0.062	0.077	0.142
Farm laborer	0.054	0.333	0.000	0.062	0.110	0.037
Forestry/fishing	0.054	0.625	0.000			

Sample = all jobs, by 2-digit occupation.

level include (a) management-related workers, (b) engineers, (c) mathematicians and computer scientists, (d) natural scientists, (e) engineering technologists, (f) service salespeople, (g) construction workers, (h) machine operators, and (i) other precision workers. In contrast, the occupations for which classical jobs are more likely to be clustered with classical jobs in dissimilar occupations include (a) public-administration workers, (b) mathematicians and computer scientists, (c) natural scientists, (d) teachers, (e) finance and business salespeople, (f) retail salespeople, (g) secretaries, (h) record keepers, and (i) assemblers.

Note that the similarities and differences in these two lists give an indication of the extent to which all modern and all classical job designs are used both within and across industries and establishments. Public administration and teaching jobs, for example, are concentrated in a narrow set of industries. Retail sales jobs are concentrated in certain types of establishments within multi-establishment firms. The tendency for classical jobs in these occupations to be concentrated with classical jobs in other dissimilar occupations helps explain the patterns in Table 7A when excluding and including controls for the type of industry. A similar argument can be made for the concentration of modern jobs for occupations such as engineers and construction workers.

In contrast, certain occupations are less likely to cluster with dissimilar occupations along both modern and classical lines, including health-related services, protective services, food services, building services, personnel services, and vehicle operators. Note that these resemble non-core activities that are likely to be found in a broad array of establishments (regardless of industry type), and thus are candidates for outsourcing (Abraham & Taylor, 1996).

5. DISCUSSION AND CONCLUSIONS

In this chapter, we presented a simple theory of job design that can be used to motivate observed trends and patterns in the empirical literature. The model is consistent with two broad approaches to job design. In the first approach, the firm uses ex ante optimization of methods. As a result, workers are given relatively narrow jobs to exploit gains from specialization and comparative advantage, and low discretion. However, ex ante optimization is not always feasible or profitable. When the firm faces greater complexity, unpredictability, or instability, it is less likely to effectively optimize production ex ante.

If so, then there is potential for the worker to learn on the job and engage in continuous improvement.

We argued that task interdependence is an important source of costs of both ex ante optimization and on-the-job learning. An alternative to ex ante optimization is continuous improvement, giving workers multitask jobs to take advantage of inter-task learning. Greater discretion complements this approach: it facilitates developing new ideas and implementing improvements. Thus, the theory is consistent with multitasking, interdependence, and discretion being positively correlated in the same job. Because the emphasis on ex ante optimization or continuous improvement depends on the firm's complexity, unpredictability, and stability, the firm's product, technology, and industry characteristics should be important factors influencing job design. Finally, there should be patterns of similar job design within firms, even more so within establishments, and also within industries.

These ideas are useful in linking the economic approach to the behavioral approach to job design, which emphasizes "intrinsic motivation" (Hackman & Lawler, 1971; Hackman & Oldham, 1976). The literature argues that multitasking and discretion may improve intrinsic motivation because the job is more intellectually challenging to the worker. Indeed, Adam Smith recognized that a cost to specialization is that workers may be bored and less motivated. The model can be interpreted as consistent with intrinsic motivation. If the marginal disutility of effort is lower when the worker performs both tasks, this yields an additional benefit to multitasking. Intrinsic motivation could be modeled by including the higher disutility of effort from specialization as one component of coordination costs of specialization compared to multitasking.

However, we purposely did not consider intrinsic motivation. Although we believe that many workers are intrinsically motivated by multitask jobs, the inter-task learning mechanism should hold regardless of any psychological effects, and is nicely complementary to the psychological explanation. The psychology story implies that multitask jobs will increase the extent to which workers are intellectually engaged in their work: thinking and curious about what they are doing. If so, this should only increase the degree of inter-task learning.

The role of skills is ambiguous in theory. Skills might reinforce the gains from specialization. However, to the extent that skills means problem-solving abilities, abstract-thinking skills, and other traits that improve the worker's learning, skills might instead reinforce continuous improvement. If so, then they would be positively associated with modern, not classical, job designs. Empirically, this is the case. This helps explain why returns to skills are

associated with technological and organizational change – they put a premium on workers making continuous improvements in production methods.

We then analyzed data on job design attributes, using reasonable proxies for our concepts of multitasking, discretion, skills, and interdependence. The results are strongly consistent with our predictions. All of the job design attributes are strongly positively correlated. There is a tendency for firms to choose either a modern or classical job design approach, but not both (at the establishment level). This is consistent with our argument that job design approaches vary with the firm's product and market characteristics. At the firm level, in contrast, there is a tendency to push job design toward extremes, choosing modern job design in some establishments and classical job design in others. This is consistent with multi-establishment firms using establishments to isolate different types of jobs (and overall organizational design emphasis on centralized, ex ante versus decentralized, continuous optimization) from each other to capture the benefits of job design while minimizing the potential downsides from doing so. At the industry level, computer usage is related to both greater use of modern jobs and greater use of classical jobs. R&D spending, in contrast, is associated only with greater use of modern jobs. This provides further evidence that job design decisions depend on the firm's product and market characteristics.

We find strong evidence that firms choose coherent job design strategies, and that the same strategy is not optimal for all organizations. The current data provide some information on characteristics of the establishment's environment that may affect this choice: larger establishments are more likely to choose modern job design, while unionized and nonprofit organizations are less likely to choose either "all classic" or "all modern" job design. There are important differences across industries in the choice of job characteristics. In future work we hope to explore this area more thoroughly to determine whether technological considerations, market structure, competition, uncertainty, or product characteristics affect the design of jobs.

NOTES

1. Inter-task learning can also occur across workers through collaboration, but with coordination costs. A more complex model might consider whether a group can learn more or less effectively than an individual. The individual does not suffer from coordination costs of getting the team to function effectively. However, a well-functioning team might learn more effectively because of the value of different priors, points of view, etc.

2. Our goal here is not to model agency costs, so we assume the simplest form. One might extend the argument to predict that worker incentives will be complementary with discretion (Holmstrom & Milgrom, 1991, 1994; Ortega, 2004). Dessein and Santos (2006) consider this possibility, and show that increasing agency costs with greater discretion may make the relationship between multitasking and interdependence nonmonotonic. Our data do not contain sufficient information on compensation policies (see footnote 3 below) to test this, so we ignore that possibility.

3. This variable in the NCS indicates that only 3.2 percent of all jobs receive incentive pay, yet jobs that include incentive-based pay account for approximately 30 percent of all jobs in the economy (Lemieux, MacLeod, & Parent, 2007). Thus the NCS definition of incentive pay clearly is an extremely narrow measure that excludes many important sources of variable or incentive pay. For this reason we do not use the NCS incentive pay measure.

4. For a detailed description of the NCS, see Pierce (1999).

5. The remaining five are: personal contacts, purpose of contacts, physical demands, work environment, and supervisory duties. We do not use these because they are not clearly linked to the choice between centralized and decentralized job design.

6. An interesting way to think about these variables is that guidelines is a form of ex ante control, useful for foreseeable contingencies, while supervision received is a form of control used for more unpredictable or idiosyncratic events.

7. Our main results are essentially unchanged even without the inclusion of this variable in the analysis.

8. When controlling for industry-fixed effects, the point estimates in Table 2 versus Table A1 do not change much, though the explained variation increases and the increase in explanatory power for each of the models is significant with a p-value <0.00001. Thus industry differences account for part of the relationship between job design attributes; they just do not account for much of the positive correlations.

9. To simplify presentation, for the remainder of the paper we use guidelines as the sole proxy for discretion. Results are very similar for supervision received. We presented results for both proxies to this point simply to illustrate similarity in the findings.

10. The standard errors in Tables 4, 5A–B and 7 were adjusted to control for intra-group correlation due to observing multiple jobs in the same establishment.

11. For sake of comparison, Appendix Table A2 contains the results when all 81 unique categories are entered separately.

12. A different response bias, in which some occupations are rated systematically higher than others even if they should not be, is already controlled for by differencing observed values for each job design attribute from the three-digit occupation-specific mean.

13. Note that each establishment is assigned its own industry classification, which may differ from that of the parent firm's. This means that some of the establishment level (across industry) variation in the first set of columns represents within-firm variance (across establishments) within large integrated firms. Consequently, when the positive correlation between the fraction of modern jobs elsewhere in the firm and the probability of a job being classical becomes insignificant (when controlling for industry-fixed effects), this may partly be due to controlling for the within-firm variance in the large integrated firms.

REFERENCES

Abraham, K. G., & Taylor, S. K. (1996). Firms' use of outside contractors: Theory and evidence. *Journal of Labor Economics*, *14*(3), 394–424.

Appelbaum, E., & Batt, R. (1994). *The new American workplace: Transforming work systems in the United States*. Ithaca, NY: ILR Press.

Autor, D. H., Katz, L. F., & Krueger, A. B. (1998). Computing inequality: Have computers changed the labor market?. *Quarterly Journal of Economics*, *113*(4), 1169–1214.

Autor, D. H., Levy, F., & Murnane, R. J. (2002). Upstairs downstairs: Computers and skills on two floors of a large bank. *Industrial and Labor Relations Review*, *55*(3), 432–447.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, *118*(4), 1279–1334.

Batt, R. (2002). Managing customer services: Human resource practices, quit rates, and sales growth. *Academy of Management Journal*, *45*(3), 587–597.

Batt, R., & Moynihan, L. (2002). The viability of alternative call centre production models. *Human Resource Management Journal*, *12*(4), 14–34.

Becker, G., & Murphy, K. (1992). The division of labor, coordination costs, and knowledge. *Quarterly Journal of Economics*, *107*(4), 1137–1160.

Bresnahan, T. F., Brynjolfsson, E., Hitt, L. M. (2002). Information technology, workplace organization and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics*, *117*(1), 339–376.

Cappelli, P., & Neumark, D. (2001). Do 'high performance' work practices improve establishment-level outcomes?. *Industrial & Labor Relations Review*, 737–775.

Caroli, E., & Van Reenen, J. (2001). Skill-biased organizational change? Evidence from a panel of British and French establishments. *Quarterly Journal of Economics*, *CXVI*(4), 1449–1492.

Cohen, S. G., & Bailey, D. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management*, *23*(3), 239–290.

Dessein, W., & Santos, T. (2006). Adaptive organizations. *Journal of Political Economy*, *114*(5), 956–995.

Gibbs, M., & Levenson, A. (2002). The economic approach to personnel research. In: S. Grossbard-Shechtman & C. Clague (Eds), *Expansion of economics: Towards a more inclusive social science*. New York: M.E. Sharpe.

Gibson, C., & Cohen, S. (2003). *Virtual teams that work*. San Francisco, CA: Jossey-Bass.

Goldin, C., & Katz, L. F. (1998). The origins of technology-skill complementarity. *Quarterly Journal of Economics*, *113*(3), 683–732.

Hackman, J. R., & Lawler, E. E., III. (1971). Employee reactions to job characteristics. *Journal of Applied Psychology*, *55*(3), 256–286.

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance*, *16*, 250–279.

Holmstrom, B., & Milgrom, P. (1991). Multi-task principal-agent analyses: Incentive contracts, asset ownership and job design. *Journal of Law, Economics and Organization*, *7*, 24–52.

Holmstrom, B., & Milgrom, P. (1994). The firm as an incentive system. *American Economic Review*, *84*, 972–991.

Huselid, M. (1995). The impact of human resource management policies on turnover, productivity, and corporate financial performance. *Academy of Management Review*, *38*(3), 635–672.

Ichniowski, C., Shaw, K. (1995). *Old dogs and new tricks: Determinants of the adoption of productivity-enhancing work practices*. Brookings Papers: Microeconomics, pp. 1–65.

Ichniowski, C., Shaw, K., & Prennushi, G. (1997). The effects of human resource management practices on productivity: A study of steel finishing lines. *American Economic Review*, *87*(3), 291–313.

Jensen, M., & Wruck, K. (1994). Science, specific knowledge, and total quality management. *Journal of Accounting and Economics*, *18*(3), 247–287.

Lawler, E. E., III., Mohrman, S. A., & Benson, G. (2001). *Organizing for high performance: Employee involvement, TQM, reengineering, and knowledge management in the Fortune 1000*. San Francisco: Jossey-Bass.

Lemieux, T., MacLeod, W. B., and Parent, D. (2007). *Performance pay and wage inequality*. National Bureau of Economic Research Working Paper no. 13128. Cambridge, MA.

Levenson, A. (2007). *The economic analysis of teams: An interdisciplinary perspective*. mimeo University of Southern California.

Levy, F., & Murnane, R. (2005). *The new division of labor: How computers are creating the next job market*. Princeton, NJ: Princeton University Press.

Lindbeck, A., & Snower, D. (2000). Multi-task learning and the reorganization of work: From Tayloristic to holistic organizations. *Journal of Labor Economics*, *18*(3), 353–376.

MacDuffie, J. P. (1995). Human resource bundles and manufacturing performance: Organizational logic and flexible production systems in the world auto industry. *Industrial and Labor Relations Review*, *48*, 197–221.

Milgrom, P., & Roberts, J. (1990). The economics of modern manufacturing: Technology, strategy, and organization. *American Economic Review*, *80*(3), 511–528.

Milgrom, P., & Roberts, J. (1995). Complementarities and fit: Strategy, structure, and organizational change in manufacturing. *Journal of Accounting and Economics*, *19*(2–3), 179–208.

Mohrman, S. A., Cohen, S. G., & Mohrman, A. M., Jr. (1995). *Designing team-based organizations: New forms for knowledge work*. San Francisco, CA: Jossey-Bass.

Morita, H. (2001). Choice of technology and labour market consequences: An explanation of U.S.-Japanese differences. *The Economic Journal*, *111*, 29–50.

Murphy, K. (1986). *Specialization and Human Capital. Ph.D thesis*. University of Chicago.

Neal, D. (1999). The complexity of job mobility among young men. *Journal of Labor Economics*, *17*(2), 237–261.

Ortega, J. (2004). *Employee discretion: Stylized facts for Europe*. Working Paper. Universidad Carlos III de Madrid, Madrid.

Osterman, P. (2000). Work reorganization in an era of restructuring. *Industrial and Labor Relations Review*, *53*(2), 179–196.

Pierce, B. (1999). Using the National Compensation Survey to predict wage rates. *Compensation and Working Conditions*, *Winter*, 8–16.

Porter, L. W., Lawler, E. E., III., & Hackman, J. R. (1975). *Behavior in organizations*. New York: McGraw Hill.

Shaw, K. (1987). Occupational change, employer change, and the transferability of skills. *Southern Economic Journal*, *53*(3), 702–719.

Smith, A. (1776). *The wealth of nations*. Reprinted by Modern Library.

Taylor, F. (1923). *The principles of scientific management*. New York: Harper.

Zoghi, C. (2002). *The distribution of decision rights within the workplace: Evidence from Canadian, Australian and UK establishments*. BLS Working Paper no. 363. Washington, DC.

Zoghi, C., & Pabilonia, S. (2004). *Which workers gain from computer use*? BLS Working Paper no. 373. Washington, DC.

APPENDIX A

Table A1. Relationships between Pairs of Job Design Attributes Controlling for Industry or Occupation.

	Controlling for Industry				Controlling for Occupation			
	Guidelines	Supervision	Skills	Interdependence	Guidelines	Supervision	Skills	Interdependence
(a) Full sample								
Multitasking	4.403	3.867	1.780	3.969	2.488	3.582	2.715	2.434
	(0.4904)	(0.5067)	(0.4473)	(0.4971)	(0.5514)	(0.5575)	(0.5358)	(0.4795)
Guidelines		3.929	1.542	3.731		2.791	2.184	2.711
		(0.5094)	(0.4267)	(0.5403)		(0.5233)	(0.5357)	(0.4953)
Supervision			1.724	3.504			1.876	3.208
			(0.4445)	(0.4842)			(0.5106)	(0.5424)
Skills				2.986				1.919
				(0.3369)				(0.3418)
(b) Non-managers								
Multitasking	4.419	3.870	1.891	3.854	2.233	2.647	2.113	3.283
	(0.4707)	(0.4878)	(0.4420)	(0.4732)	(0.4217)	(0.5331)	(0.5230)	(0.5254)
Guidelines		3.872	1.676	3.640		2.847	2.426	3.430
		(0.4869)	(0.4344)	(0.5213)		(0.4965)	(0.5351)	(0.5524)
Supervision			1.807	3.443			2.549	3.168
			(0.4381)	(0.4665)			(0.5061)	(0.5175)
Skills				3.072				2.385
				(0.3402)				(0.5377)
(c) Managers only								
Multitasking	4.273	4.021	3.503	3.595	4.257	3.885	3.444	2.675
	(0.4473)	(0.4583)	(0.4320)	(0.4906)	(0.4328)	(0.4330)	(0.4188)	(0.4001)
Guidelines		3.070	2.200	2.942		4.541	2.752	3.994
		(0.3709)	(0.2998)	(0.4318)		(0.4590)	(0.3309)	(0.5352)
Supervision			2.883	3.502			2.797	3.415
			(0.3843)	(0.4618)			(0.3640)	(0.4433)
Skills				2.903				3.011
				(0.4182)				(0.3970)

Relationships between factors are coefficients from fixed-effect ordered logits; each cell represents a separate logit. Rows are dependent variables; columns are independent variables. Pseudo-R^2 are in parentheses. The 1990 US Census 3-digit industry and occupation codes were used to define the industry and occupation

Table A2. Effect of Distribution of Other Jobs' Characteristics on Probability of HHHH or LLLL Job Design.

	Pr(LLLL)		Pr(LLLL)		Pr(HHHH)		Pr(HHHH)	
Industry Controls	No		Yes		No		Yes	
% Other Jobs with	Establishment	Firm	Establishment	Firm	Establishment	Firm	Establishment	Firm
LLLL	**2.930***	**3.034***	**2.039***	**1.982***	**−0.7171***	0.3645	**−0.8802***	0.2058
MLLL	**1.074***	**1.625***	**0.9392***	**1.254***	**−0.7024****	**−1.345***	**−0.8194****	**−1.223****
LMLL	**1.844***	0.2090	**1.738***	−0.1919	−0.2427	−0.9928***	−0.4628	−1.153***
LLML	0.9338***	0.1835	**1.234***	0.5774	−1.234***	−0.1717	−0.9048	0.0405
LLLM	0.9570***	1.160	**1.169****	**2.239***	0.4330	0.6796	0.2028	0.6820
LLMM	**1.122***	0.3567	**1.133***	−0.0145	**−1.024***	**−1.003***	**−0.8834***	−0.7334***
LMLM	0.9415	1.622	0.7171	1.029	−1.824***	0.5012	**−2.455****	−0.1444
LMML	**2.303***	−1.604	**2.426***	−1.600	−1.629***	−0.2104	−1.948***	−1.410
MLLM	0.7638***	−0.6311	**0.8967****	−0.7383	−0.6879	0.4281	−0.4200	0.8707***
MLML	0.3937	0.9660	0.6890	1.262	**−1.576****	−1.614***	**−1.537****	−0.7724
MMLL	**1.591***	−0.1566	**1.351***	−0.2677	0.4982	0.1221	0.1671	−0.4334
LMMM	**0.7535***	**0.6972****	**0.7321****	0.4653	−0.3561	**−1.106***	−0.2394	−0.8151***
MLMM	**0.6803***	**−0.7866****	**0.6667***	**−0.7356****	**−0.9819***	**−1.174***	**−0.7175***	−0.5592***
MMLM	0.3717	0.7183	0.2871	0.2194	−0.3650	−0.8727	−0.4987	**−1.393****
MMML	0.0673	**1.344****	0.2761	0.6067	**−1.610***	0.4538	**−1.978***	0.6151
LLLH	4.776	−110.3	−0.9032	−150.8	1.185	2.907	1.769	0.1019
LLHL	3.561	6.108	2.391	1.221		−4.927		−46.84
LHLL	1.657	4.927	1.529	5.169	−0.8927	−0.6115	−1.127	−1.811
HLLL								
LLHH	−0.6134	6.032	−4.319	2.270	−7.676	7.197	−8.787	8.989
LHHL, HLLH, HLHL have no observations								
HHLL						−38.19		−51.80
LHHH	−0.3998	−17.99	−0.1042	−18.45	−5.380	−15.898	−5.326	−17.13
HLHH	2.028	1.647	2.523	2.622	0.1989	−1.965	2.308	2.711

Table A2. (*Continued*)

	Pr(LLLL)		Pr(LLLL)		Pr(HHHH)		Pr(HHHH)	
Industry Controls	No		Yes		No		Yes	
% Other Jobs with	Establishment	Firm	Establishment	Firm	Establishment	Firm	Establishment	Firm
HHLH	−1.508	11.60	−1.463	18.17***	−3.642	2.226	−5.912	0.3935
HHHL	0.6190	−5.713	−0.4054	−19.07	−0.3951	16.01		12.89
HMLL	7.553	−20.25	9.296	−13.67	−7.700	0.0131	−8.809	−0.7919
HLLM	−3.164		−9.150		1.100		2.320	
HLML		16.31		17.72	1.672	4.573	0.0613	1.326
LLMH	−1.218	1.124	−1.668	1.138	1.610	0.4258	−1.272	−2.921
LLHM	−2.698	−6.065	−2.379	−4.521	**6.582***	5.145	**5.594****	1.868
LMLH								
LMHL	−6.866	7.909	−8.165	3.076	−0.5411	−0.6187	−1.655	−1.677
LHLM	−1.098	−7.354	−1.007	−10.33	−5.564	−1.287	−8.235	−4.626
LHML	0.2737	−1.477	1.737	−0.4743	1.080	−2.689	0.2264	−4.215
MLLH						3.936		3.187
MLHL		−17.15		−7.612	−10.22	1.975	−11.02	4.926
MHLL	0.4121	−0.3604	0.3802	0.8369	−0.5419	**2.338****	−1.262	1.602
LHMM	1.182	1.492	1.271	0.7813	1.262	0.1630	0.9830	−0.7765
LMHM	−0.0570	−1.117	0.3534	−1.771	**3.639***	−0.6835	**4.194***	−2.147
MLMH	**−8.101****	**4.030****	**−8.666****	**4.437****	−5.267***	−0.2540	−5.238***	0.8430
LMMH	−2.468	1.155	−2.037	2.536***	−0.2871	−4.371	−0.3278	−4.442
MLHM	−2.994	−1.163	**−5.063****	−5.333	−0.6571	−2.246	−0.7093	−1.076
MMLH	−1.838	**2.712*****	−3.646	2.362	0.0641	−2.400	−1.248	−3.298
MMHL	−1.082	0.4636	−1.732	−0.9096	−0.2973	1.884	0.4413	2.465
MHLM	1.829	−0.6053	2.773***	0.7817	−0.4844	0.9438	−1.331	0.3649
MHML	0.5310	1.622	1.035	1.637	0.3212	0.8711	0.2347	0.7884
HMML	−1.089	−0.5937	−3.585	−0.2369	−3.512	1.572	−3.183	−0.0244
HLMM	−0.5964	−3.946	0.0959	−4.768	−1.776	−0.0359	−2.187	0.2914
HMLM	3.021	1.560	3.412	3.604	−25.77***	−0.7417	−25.69***	−3.794

MMMM = base case								
LHHM	3.748	8.707	2.430	2.675	−5.210	6.151	−5.910	−4.125
LHMH	7.090	−2.587	2.396	−11.98	−9.246	5.130	−7.530	11.95
LMHH	−2.878	2.628	−6.881	0.1668	−7.683	−7.829	−4.645	−4.754
MHLH		1.380		−2.372				
MLHH	2.908***	0.1066	2.473	−3.005	0.3599	0.9696	0.3063	0.3924
MHHL	−0.3631	−6.298	−1.121	−5.609	1.926	1.094	1.941	1.828
HLMH	0.5681	0.3618	1.290	1.502	0.7120	−4.041	1.926	0.6389
HLHM	−6.722	2.506	−5.985	−0.5327	3.174	0.1946	2.531	1.259
HMLH								
HMHL	−2.667	3.272***	−1.967	**5.469****	1.292	1.999	1.058	1.548
HHLM	3.278	3.184	3.177	1.191	−1.645	7.408	−1.038	6.143
HHML	10.06***	−8.923	6.349	−11.11	8.463	−66.81	6.946	−61.94
HMMM	−0.1239	0.4239	−0.4195	−0.0931	**0.9412***	−0.0846	**0.7902***	−0.2015
MHMM	**−0.9904***	0.4371	−0.7566***	0.9236***	0.2081	.2558	−0.0889	−0.2044
MMHM	**−3.332***	−0.6602	**−3.068***	−0.1621	0.5723	**−1.473***	**0.9796***	−0.8266
MMMH	−0.2135	**−0.8186****	−0.7285***	0.3747	−0.2489	−0.2248	0.0724	
HHMM	−0.5866	0.3929	−0.8696***	−0.7068	**1.394***	**1.044****	**1.220***	0.3106
HMMH	−1.334	−0.0209	−1.449	0.2831	0.2818	0.2708	0.0974	0.0132
HMHM	−0.1646	0.2349	−1.068	−0.2462	1.596*	−1.080	1.211*	−1.862**
MHMH	−2.012***	−2.688***	−1.823***	**−3.515****	0.7688	0.7720	0.7921	0.5786
MHHM	**−1.432****	−1.338***	**−1.627****	**−1.857****	**−0.8736****	**−1.313****	−0.4818	−0.3751
MMHH	−0.2920	−0.2982	−0.5473	−0.5960	**1.743***	**1.034***	**1.537***	0.5108
MHHH	**−0.9279****	−0.0668	**−1.147****	0.1604	**0.7060****	0.1495	**0.7671***	0.6435
HMHH	−0.1202	**1.402***	−0.8495	0.5875	**2.171***	0.6445***	**2.113***	0.5864
HHMH	**−1.212****	−0.1656	−1.075***	−0.5682	−0.0479	0.6851***	−0.1722	0.5475
HHHM	−0.5029	−0.5579	−0.4967	−0.5042	**1.640***	**2.222***	**2.082***	−0.6012
HHHH	−1.076*	0.7060**	−1.252*	0.3912	3.054*	1.640*	2.483*	1.101*
R^2	0.1029		0.1225		0.1270		0.1389	
N	41,164		40,028		41,323		40,472	

*p-value < 0.01.
**p-value < 0.05.
***p-value < 0.10.

APPENDIX B. DISCUSSION OF MODEL

Proof that Multitasking is Preferred to Specialization for Some Range of k *and* s

From (1) and (3), multitasking is preferred to specialization if:

$$2\alpha^{\alpha}\left(\frac{s+k}{1+\alpha}\right)^{1+\alpha} > s^{1+\alpha} - C$$

For simplicity, assume that $C = 0$; if $C > 0$, multitasking is even more likely to be preferred. The condition above can then be rewritten as:

$$\left(\frac{s+k}{s}\right)^{1+\alpha} > \left(\frac{\alpha}{2}\right)\left(\frac{1+\alpha}{\alpha}\right)^{1+\alpha}$$

For fixed α and s, some k^* exists for which this expression holds for all $k > k^*$, since the left side is increasing in k (and similar logic would apply for any fixed $C > 0$). Setting both sides equal and solving yields:

$$k^* = s\left[\left(\frac{\alpha}{2}\right)^{1/1+\alpha}\left(\frac{1+\alpha}{\alpha}\right) - 1\right]$$

Proof of Equation (5)

$Q_{\text{multitask|centralization}} = \max_{\tau}[E(Q)] =$ expected output with τ chosen over the entire distribution of the unknown state of the world. $Q_{\text{multitask|discretion}} = \max_{\tau}[Q \mid \text{state of the world}]$. The τ chosen to maximize expected output can result in actual output no better than when the state of the world is known. Since this logic applies for any given state of the world, it also applies unconditional on the state of the world, as in Eq. (4). Finally, if there were agency costs associated with discretion, the worker would be given discretion only if the benefits outweighed those agency costs.

IS SENIORITY-BASED PAY USED AS A MOTIVATIONAL DEVICE? EVIDENCE FROM PLANT-LEVEL DATA

Alberto Bayo-Moriones, Jose E. Galdon-Sanchez and Maia Güell

ABSTRACT

In this chapter we use data from industrial plants to find out whether seniority-based pay is used as a motivational device for production workers. Alternatively, seniority-based pay could simply be a wage-setting rule independent of incentives. Unlike previous papers, we use a direct measure of seniority-based pay as well as measures of monitoring devices and explicit incentives. We find that those firms that base their wages partly on seniority are less likely to offer explicit incentives. They are also less likely to invest in monitoring devices. We also discover that these companies are more likely to engage in other human resource management policies, which result in long employment relationships. Overall these results suggest that seniority-based pay is indeed used as a motivational device.

Jobs, Training and Worker Well-Being
Research in Labor Economics, Volume 30, 155–187
Copyright © 2010 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 0147-9121/doi:10.1108/S0147-9121(2010)0000030008

1. INTRODUCTION

The use of incentive schemes is an important instrument to motivate workers in the hands of firms. However, not all firms and not all types of production processes require the same kind of incentives to enhance motivation and, therefore, labor productivity.

There are two basic types of incentives that firms may offer in order to motivate workers: explicit and implicit incentives. Explicit incentives are a direct way to stimulate workers basing their pay on productivity, for instance through a contract that allows for piece rates. This type of incentive is appropriate when the individual output is easy to observe and quantify. When this is not the case, implicit incentives not directly connected to productivity are an alternative way to motivate workers.

In this chapter, we concentrate on implicit incentives. More specifically, we focus on a particular wage contract that can mostly serve as a deferred compensation scheme and, therefore, as a motivational device: the seniority-based pay contract. In a context in which workers can be fired for disciplinary reasons, workers who receive higher wages at higher seniority levels would be motivated to work harder in order to avoid getting fired and thus would obtain higher future wages.

In general, a seniority-based pay contract implies a contract in which, at some point in the worker's life cycle, there is a discrepancy between the spot wage and the spot value of the worker's marginal product. Such contracts can act as motivators. The reason is that workers are paid below their productivity during the first few years of their contract, while their wage is above their productivity in the final stage of their career at the firm. If workers do not shirk, they will be allowed to stay at the firm and will be able to recuperate their initial losses. If they shirk, however, they run the risk of being caught and dismissed, and therefore, they jeopardize the chance to recover the wages that the firm owed to them, in the last years of their contract. According to Lazear (1979), workers and firms enter into these long-term implicit contracts to discourage shirking and malfeasance by shifting compensation to the end of the contract.

But seniority-based pay could be a wage rule completely independent of incentives. For instance, it could be a practice due to cultural or social norms, or to the presence of unions, to induce self-selection of some type of workers in the firm or to insure risk-averse workers who are uncertain about their productivity.[1]

According to Lazear, when seniority-based pay is used as a motivational device, seniority-based pay and explicit incentives should be negatively

correlated. Similarly, since firms may undertake monitoring activities if input is easy to observe and quantify, when seniority-based pay is used as a motivational device, it should be negatively correlated with the use of monitoring devices.

However, when seniority-based pay is simply a wage rule, the existing theories are silent about its relationship with the provision of explicit incentives (or monitoring); that is, there is no theoretical reasoning that predicts any particular relationship between these two sets of practices. Therefore, in this case, we should not expect these two sets of policies to be related in any systematic way.

Our empirical strategy, to identify whether firms that base wages on seniority do it for motivational reasons, will be based on Lazear's theory. Additionally, we investigate the relationship between seniority-based pay and other firm practices. The reason is that, as we should expect, firms that use seniority-based pay for motivational reasons should be more likely to undertake practices that involve long-term employment relationships than firms that provide explicit incentives or use monitoring devices. Finally, we also analyze the relationship between seniority-based pay and training provided by the firm.

Incentive theories have been difficult to test empirically due to the lack of available data (see Lazear, 1979). This problem is perhaps even more serious with regard to implicit rather than explicit incentives.[2]

In the past, there have been several attempts to test the theory presented by Lazear (1979) with the available datasets (see, e.g., Lazear & Moore, 1984; Hutchens, 1987; Barth, 1997, among others).[3] Most of these studies tested the predictions of the theory in terms of worker's earnings and productivity, dismissals and tenure, or the incidence of mandatory retirement or workers' pensions. A common characteristic of the existing labor literature is that it uses survey data at the worker level, assuming the existence of a relationship between wages and seniority. However, to our knowledge, there is not yet a study that has used a direct measure of the existence of a seniority-based pay contract at the firm level. In this chapter, we test empirically the theory of implicit incentives using new data, which allows us to directly observe whether firms decide to set workers' wages according to seniority or not.[4]

We use a unique plant-level dataset that contains direct information on several firms' personnel practices for 734 industrial establishments in Spain. All surveyed establishments are involved in production processes within the manufacturing sector. Regarding personnel practices, the survey refers to the blue-collar workers in each plant (i.e, workers involved directly in

production). Overall, we obtain very homogeneous data for every surveyed plant. At the same time, a wide scope of different firms within the manufacturing sector is included in the survey.

The main feature of the dataset is that it refers to firms rather than individuals and that it contains a considerable number of firms. This allows us to measure the presence of seniority-based pay from a different perspective than the one traditionally used in the empirical literature of tenure and wages, which concentrates on worker-level data. Similarly, our data allows us to obtain direct measures of monitoring devices, as well as other measures of explicit incentives practices. Moreover, the use of plant-level data allows us to get a better understanding of the role that firm and job characteristics play in the diffusion of deferred payment schemes. This is a question that has been scarcely addressed in the literature that uses data at the plant level.

Spain is an interesting case to analyze seniority-based pay because, in this country, retirement is mandatory for all workers once they reach the age of 65. Therefore, all establishments in our sample are subject to this mandate. According to Lazear (1979), jobs with delayed payment contracts should be characterized by mandatory retirement. The institution establishes a termination date after which the worker is not entitled to continue receiving a wage that is greater than her productivity. Moreover, production workers in Spanish manufacturing firms do not have absolute employment security. If they do not perform well in their job, they can be fired for disciplinary reasons (see Galdon-Sanchez & Güell, 2003). In this context, seniority-based pay can become an optimal contract.

Our results provide empirical support for the theories that are behind the deferred wage schemes as motivational devices. We find that firms that offer seniority-based pay are less likely to offer explicit incentives. They are also less likely to invest in monitoring devices. We also find that firms that offer seniority-based pay rather than explicit incentives are more likely to engage in other personnel practices that imply long employment relationships. Finally, since seniority-based pay could be related to other personnel practices, especially training, we also analyze whether this is the case in our dataset.

The rest of the chapter is organized as follows. In the next section we review the related literature. Section 3 is devoted to the description of the survey from which we obtained the data used to perform our exercise. In Section 4 we undertake our empirical analysis. We define all the variables used in our exercise and proceed to the descriptive analysis of such variables. The results appear in Section 5, which is followed by the conclusion.

2. LITERATURE REVIEW

In this section, we review all the literature related to our study. First of all, we review the theory that considers the seniority-based pay contract as a motivational device, as well as the existing empirical evidence on this theory. Second, we review other theories that consider seniority-based pay contracts as independent of incentives. We also review some of the literature on the correlation of seniority and wages due to the provision of human capital.

2.1. Seniority Pay as a Motivator Device

2.1.1. Theory

Lazear (1979, 1981) offers an explanation for seniority-based pay founded on motivational issues. Seniority-based pay can be used to align the interests of the worker with those of the company and, therefore, lead to greater levels of effort from employees.

If the firm offers a wage profile in which, at every moment in her career, the employee is paid just for her productivity, when the worker is close to retirement she will be likely to shirk. This will happen because, in that case, the worker has nothing to lose since she is not going to be employed anyway any more. The same would apply at any other moment during her working life at the firm, as long as the costs of finding a new job at that moment are small.

However, this behavior would not take place if the worker had to face a penalty for not putting enough effort. For example, the employee had to pay an up-front fee, which would be returned at the moment of retirement from the firm. The fee would only be returned if the employee had not shirked during her career in the company. But, if she shirked, she would be fired, not ending her career at the firm and not getting back the up-front fee.

One way of generating this penalty is linking wages to seniority, paying the employee below her productivity at the beginning of her career and above her productivity at the end of it. This linkage implies a steeper association between wages and seniority than that between productivity and seniority. It also implies that seniority pay becomes a deferred compensation scheme. By allowing initial wages to be paid at the end of the employee's career, the firm discourages its workforce from engaging in any inappropriate behavior. This increases both the value that the employee can be expected to contribute to the firm and the total amount of wages that this worker receives throughout her career in the firm. This motivation

mechanism can only work if it goes together with the threat of dismissal in case of poor job performance. If firing is not possible due to certain reasons (such as legal constraints), seniority pay cannot act as a motivator device.[5]

There are several implications from the theory in Lazear (1979). The first one is that, if used as a motivational device, seniority-based pay is unnecessary when there are other mechanisms that prevent employees from shirking. If effort can be easily observed by the firm, for example because many resources are devoted to monitoring, the motivation role of seniority pay does not make any sense. The same applies when worker's output is easy to observe by the firm. In this case the firm can use explicit incentive mechanisms, such as payment by results, to encourage worker's effort. As a consequence, a negative relationship between the use of seniority pay and both the degree of monitoring and the existence of explicit incentives should be expected.

The second implication of Lazear's theory is that those jobs characterized by seniority-based pay should have higher wage growth rates than productivity growth rates. If seniority pay is used as a motivating device, it must be a deferred compensation scheme. For that reason, it is a long-term implicit incentive that involves a promise from the firm to the employee. The employee will accept this implicit contracting if she trusts that the firm will not renege on its promise. This will happen when the company has a solid and established reputation as employer, which is more likely to be associated with firm features such as size or age.

Other implications of this theory, which are not that relevant for the purpose of this chapter, are as follows: (1) pensions, which discourage shirking until the end of a labor relation, are more common in situations that include implicit incentives; (2) firms with jobs that included implicit incentives should implement mandatory retirement in order to fix a termination date, after which wages cannot grow beyond productivity; and (3) long-tenured workers are more likely to have jobs that offer mandatory retirement and pensions.

2.1.2. Empirical Evidence

Empirical tests of Lazear's theory have been performed from very different perspectives. Since the implications that derive from this model are applicable to quite different research areas, it is possible to analyze its validity with different empirical approaches.

Some authors have concentrated their efforts on testing the prediction on worker's wage growth, that is, wages rise more rapidly than productivity

(see, e.g., Medoff & Abraham, 1980; Lazear & Moore, 1984; Spitz, 1991; Lazear, 2000). Other authors have studied the implications of the theory with regard to mandatory retirement and earnings. Examples of these are the original paper by Lazear (1979) and the paper by Clark and Ogawa (1992), which tests the theory in Japan. Alternative approaches have studied the implications of the theory for dismissals and tenure (see, e.g., Idson & Valletta, 1996).

There are also other empirical articles that, just like ours, focus more directly on the motivational nature of seniority pay. They test the hypothesis that seniority-based pay will be applied in circumstances of agency problems, provided that the link between wages and seniority is used as a means to motivate workers. This problem does not happen when workers are self-employed, that is, when they own the firm in which they work.

If seniority pay acts as a motivator to workers, the wage-seniority slopes found in self-owned companies ought to be less pronounced than those found in other types of firms. Lazear and Moore (1984) find empirical evidence to support this argument since, in the case of self-employed workers, the present value of the lifetime income earned by an employee increases less with the slope of the age-earnings profile.

Hutchens (1987) focuses on the relationship between seniority pay and monitoring, which depends on how repetitive are tasks in a job. Then, the author analyzes jobs according to the predicted characteristics of Lazear's theory (long tenure, pensions, mandatory retirement, etc.) taking into account the degree of monitoring that workers are subject to.

Another paper that relates to our work is Barth (1997). Based on a sample of Norwegian workers, the author reports that employees paid on a piece-rate basis do not make any profit, in terms of wages, from staying with the same company over a long period of time. The author came to this conclusion by estimating a wage regression (controlling for worker seniority) and including a variable that captures the presence of piece rates along with an interaction term between piece rate and seniority.

A common characteristic of all these papers is that they use survey data at the worker level, inferring the existence of a relationship between wages and seniority from the analysis of worker's wage and tenure data. Our chapter, however, measures the existence of seniority-based pay directly at the firm level. In addition, we evaluate both monitoring devices and explicit incentives as well. Finally, we also have information on other personnel practices that have important implications for the worker's tenure and thus should be related to seniority-based payments.

2.2. Other Theories

Apart from Lazear's explanation, firms can offer seniority-based pay contracts for other reasons independent of incentives (see Hutchens, 1989). One of these alternative explanations of upward sloping wage profiles is provided by the models of self-selection (see Salop & Salop, 1976). According to these models, this increasing wage path will attract those workers who intend to stay with the company throughout their professional careers. This has a positive impact on the firm due to the reduction in staff turnover costs.

Seniority pay also appears as a delayed payment scheme if it is used as a mechanism to reassure risk-averse workers who are insecure about their productivity (see Harris & Holmstrom, 1982). Social norms or cultural reasons could also explain the presence of such contracts. In these cases the origin of seniority-based pay could be the preference of workers for rising earnings-seniority profiles over decreasing or flat profiles (Loewenstein & Sicherman, 1991; Frank & Hutchens, 1993) or the desire to keep status consistency inside the company (Baron & Kreps, 1999).

Another possible explanation for seniority-based pay has to do with the role of unions in the determination of working conditions (Freeman & Medoff, 1984). In those workplaces undergoing collective bargaining, unions will tend to favor the situation of their members with more power, who usually are those with more seniority in the firm.

Finally, others have explained the existence of a positive correlation between wages and tenure with different approaches. The most popular one is provided by the human capital theory (see Becker, 1964; Mincer, 1974; Felli & Harris, 1996).[6] According to this theory, the existence of specific human capital increases worker productivity with tenure in the firm. Wages reflect such productivity gains and, as a consequence, seniority has a positive influence on wages through its effect on productivity. However, the existing empirical evidence is not unambiguously consistent with the specific human capital theory. While a part of the literature has found that wage increases due to seniority have their origin in productivity increases (see, e.g., Brown, 1989; Hellerstein & Neumark, 1995), other part of the literature provides empirical evidence that shows the seeds of doubt in relation to the validity of the predictions of the specific human capital theory and suggests the possible validity of the aforementioned explanations (see Medoff & Abraham, 1980; Kotlikoff & Gohkale, 1992; Levine, 1993; Flabbi & Ichino, 2001; among others).

As it is evident, the most distinguishing feature of Lazear's theory is that it connects seniority-based pay to other practices such as explicit incentives

or monitoring. Nevertheless, these practices do not play any role in all these other theories or possible explanations for the existence of seniority-based pay contracts. Our empirical strategy will be exactly based on this feature.

3. SURVEY'S DESCRIPTION AND DATA

In this chapter, we use a unique dataset that contains plant-level information on several firms' policies. All surveyed establishments are involved in production processes within the manufacturing sector. Overall, we obtain very homogeneous data for every surveyed plant. At the same time, a wide scope of different firms within the manufacturing sector is included in the survey. Next, we describe the characteristics of the survey and concentrate on the variables that we are going to use in our analysis.

The data were carefully collected in 1998 in the context of a wider research project on human resource management and operations management in Spain's manufacturing industry. All answers to the questionnaires refer to 1997. The concept of manufacturing industry is clearly defined in the Spanish National Classification of Economic Activity (Clasificacion Nacional de Actividades Economicas; CNAE),[7] which includes all the manufacturing industries with the exception of oil refining industry and that of the treatment of nuclear fuel.

The manufacturing industry was therefore chosen as the research focus for several reasons. First of all, it is a sector in which heterogeneity is limited compared to, for example, other sectors such as services. Second, manufacturing is an industry with a considerable weight in the economy of Spain. This allows us to draw more general conclusions, which would be applicable to a wider range of firms. Finally, choosing a wide scope of activities within the manufacturing sector allows us to obtain fairly general conclusions, while we avoid the problems of datasets that are too general and heterogeneous (see Ichniowski & Shaw, 2003).

When designing the survey, we decided that information should be collected at the plant level. In the manufacturing sector the plant is the basic business unit, which has strategic importance for the implementation of the practices under study. These practices are adopted in the plant, and therefore, it is at this level that problems arise and where results must be analyzed. Moreover, the answers to the different questions raised are expected to be more reliable when taken at the plant level, since knowledge of these issues is greater at this level, even if it is only for reasons of greater proximity to the matters addressed in the survey.[8]

Another aspect of the research scope to be defined was the size of the establishments to be analyzed. The industrial plants included in our sample employ 50 or more workers. Other similar studies established this same limit (see, for instance, Osterman, 1994), which, in our case, serves to cover a wide spectrum of the population employed in the Spanish industry. Moreover, it simplifies the fieldwork, since there are more reliable directories of firms' population for this group.

In order to carry out the investigation, the members of the research group, together with the firm in charge of the fieldwork, designed a questionnaire, after a close examination of the international literature related to the project content. The preliminary survey was tested in nine plants. After the pilot test, the questionnaire was modified in several ways before arriving to its final version. The questionnaire was divided into the following parts: general characteristics of the establishment, technology and quality management, human resource management, work organization, relations with customers and suppliers, and information on the firm.

Regarding personnel practices, the survey refers to blue-collar workers in each plant, that is, workers directly involved in the production process. The fact that we refer to a specific group of workers could create problems, as far as generalization of results to other professions is concerned. However, limiting the occupation under study makes comparisons easier, since there are possibly several internal labor markets with substantial differences between them within a company.

The information was gathered by interviewing the plant manager or the operations or human resources manager in the plant. A personal interview was chosen as the method of collecting information because it gives a higher response rate.

The reference universe, that is, manufacturing plants with at least 50 workers,[9] was formed by 6,013 units. The aim was to obtain a sample of 1,000 units, stratified according to sector and size. The larger-size stratum was represented at 50 percent in the sample design. For the two remaining size strata, a fixed number of 30 interviews were allocated to each sector; the rest of the interviews being allocated proportionally across sectors. The sample allocated to each of the strata within a sector was also distributed proportionally. A random selection of plants was taken from each stratum for interview. After making 3,246 telephone calls to make the necessary appointments, 965 valid interviews were conducted. The sample size corresponds to 16 percent of the population. Table A1 in the Appendix displays the ratio sample to population by firm size and sector.

For the purpose of this chapter we analyze a final sample of 734 plants, those for which none of the variables have missing values.[10] For this type of data, this is a rather large sample size.

4. EMPIRICAL ANALYSIS

Our goal is to understand if there is an incentive motive when a firm offers seniority-based contracts to their employees. A crucial aspect to bear in mind is that production workers in Spanish manufacturing firms do not have absolute employment security. If they do not perform well in their job, they can be fired for disciplinary reasons (see Galdon-Sanchez & Güell, 2003).[11] In this context, seniority-based pay could become an optimal motivational contract.

In order to determine if seniority-based pay is used as a motivational device rather than a wage rule independent of incentives,[12] we proceed as follows. We first analyze the relationship among seniority-based pay, monitoring devices, and explicit incentives. When seniority-based pay is used as a motivational device, then seniority-based pay and explicit incentives (or monitoring) should be negatively correlated (see Lazear, 1979). On the other hand, when seniority-based pay is simply a wage rule, there is no theoretical reason (or any economic mechanism) that predicts any particular relationship between seniority-based pay and explicit incentives. In other words, in this case we should not expect these two policies to be related in any systematic way. More particularly, we should not observe any sizeable negative correlation.

Once we establish that seniority-based pay is negatively correlated with explicit incentives and monitoring, in a second step we analyze other practices that could be potentially important to the firm when deciding to choose seniority-based pay to motivate its workers. We specifically consider other personnel practices that favor long-term employment relationships. These practices make the firm's commitment to pay high future wages credible and, therefore, are complementary measures to implicit incentives. They provide further evidence that seniority-based pay could be used as an incentive device.

As mentioned earlier, wages can be correlated with worker's tenure for reasons other than those related to incentives. The most obvious alternative is the existence of training. Therefore, in a third step we will analyze the relationship between seniority-based pay and training policies.

Different personnel practices are usually chosen simultaneously by a firm, generating "systems" or "bundles" of practices. There are theoretical

foundations that explain the complementarities of different policies (see, e.g., Holmstrom & Milgrom, 1994). We are aware of the possible endogeneity problems of including different personnel practices as independent regressors when estimating the probability that firms use seniority-based pay schemes. However, in the present context, and, more precisely, due to this multidimensional nature of the firm's practices, it is very difficult to find instruments. Therefore, we carefully interpret our results as bivariate relationships among different personnel practices.

In the next two subsections we first describe the variables used in our exercise (Section 4.1) and, second, take on a descriptive analysis of those variables (Section 4.2).

4.1. Variables

The survey contains information on the two most important factors that are taken into account when setting the fixed part of blue-collar workers wages. The survey makes a clear distinction between the fixed part and the variable part of worker's remuneration. There are five possible factors that may determine the fixed part of wages. These include seniority, worker characteristics (skills, efficiency, evaluation from a supervisor) and job characteristics. Using the information gathered from the survey, we construct two variables that will be the main dependent variables in our exercises. The first one, *Seniority-Based Pay (incidence)*, captures whether firms use seniority-based pay or not. Among firms that use seniority-based pay, there may be differences in the degree to which such practice is being used. The second variable, *Seniority-Based Pay (intensity)* captures the different degrees to which firms may use seniority-based pay.

More specifically, the variable *Seniority-Based Pay (incidence)* takes value one when firms base wages partly on seniority, that is, if seniority was mentioned either as the most important or second most important factor when setting wages, and zero otherwise. We also constructed *Seniority-Based Pay (intensity)*, which takes value two when seniority was said to be the most important factor to set wages, value one when it was mentioned as the second most important factor, and value zero in the remaining cases. These variables directly capture the idea of seniority-based pay contracts. We find that a substantial fraction of firms followed this policy: around 30 percent of firms pay partly according to seniority. Among these, 30 percent say that seniority is the most important criteria used when setting wages, while for the remaining 70 percent it is the second most important criteria. These figures are empirically relevant to conduct our exercise.[13]

In the survey, firms were asked whether they offer incentive payments to their blue-collar workers. These included payments that are based on productivity, quality, plant-level, or firm's results. This type of incentives corresponds to the explicit incentives mentioned earlier. Using this information, we construct two variables that capture explicit incentives in a way similar to that of the two seniority-based pay variables. The first one, *Explicit Incentives (incidence)*, captures whether firms use explicit incentives or not. The second variable, *Explicit Incentives (intensity)* captures the different degrees to which firms may use explicit incentives. In particular, we define the variable *Explicit Incentives (incidence)*, to which we assigned value one when firms answered affirmatively to this question and zero when the answer was negative. As Table 1 shows, around 62 percent of firms offer some explicit incentives to their workers. We also create the variable *Explicit Incentives (intensity)*, which registers the percentage of worker's earnings that such incentives represent. On average, in our sample, this accounts for 10 percent of wages.

We repeat all our analyses using another measure of explicit incentives that focuses on firms that only offer *individual* explicit incentives, that is, using variables *Explicit Individual Incentives (incidence)* and *Explicit Individual Incentives (intensity)*, respectively (see Table 1). As will be shown, the results of the chapter are generally robust to these alternative measures of explicit incentives.

The survey also contains information on the degree of supervision and control under which manual workers perform their duties at the plant. The answers are in a scale of one to five, where one is equivalent to no supervision at all, and five is equivalent to close supervision. Using this information we construct the variable *Monitoring (incidence)*, to which value one is assigned when the degree of control is sufficiently high (i.e., values four and five as the answers to this question) and zero otherwise. In our sample, around 40 percent of firms spend resources in supervising their workers according to this variable. Similarly, we also define the variable *Monitoring (intensity)*, which takes values one to five.

We then turn to look at factors other than incentives that could also be behind the determination of seniority-based pay schemes. In our empirical analysis, it is important to control for these factors.

- Sector

  Our dataset includes information on the sector to which the plant's activity belongs (at a three-digit level). The sector indicators capture the nature of the production technology. This is crucial to determine the ease to monitor effort (see Hutchens, 1987). According to the

Table 1. Variable Definitions and Descriptive Statistics.

Variable	Definition[a]	Mean	SD
HRM practices incidence			
Seniority-Based Pay (incidence)	1 = wages partly based on seniority, 0 = otherwise	0.287	(0.452)
Explicit Incentives (incidence)	1 = explicit incentives provided, 0 = otherwise	0.619	(0.485)
Explicit Individual Incentives (incidence)	1 = explicit individual incentives provided, 0 = otherwise	0.519	(0.500)
Monitoring (incidence)	1 = workers subject to high supervision, 0 = otherwise	0.396	(0.488)
HRM practices intensity			
Seniority-Based Pay (intensity)	2 = seniority as most important factor in setting wages, 1 = seniority as second most important factor, 0 = otherwise	0.372	(0.636)
Explicit Incentives (intensity)	Percentage of earnings that correspond to incentive pay	10.17	(11.01)
Explicit Individual Incentives (intensity)	Percentage of earnings that correspond to individual incentive pay	9.158	(11.128)
Monitoring (intensity)	Level of supervision: 1 = no supervision at all, 2 = hardly any, 3 = moderate , 4 = high, 5 = very high	3.36	(0.7)
Firm characteristics			
Old	1 = plant founded before 1980, 0 = otherwise	0.738	(0.439)
StateShare	1 = state owns a share of the firm, 0 = otherwise	0.034	(0.181)
Multinational	1 = firm belongs to multinational group, 0 = otherwise	0.287	(0.452)
Large	1 = firm with more than 500 workers, 0 = otherwise	0.107	(0.309)
Union	1 = unionization of workers above 60%, 0 = otherwise	0.318	(0.467)
Wage Level Above	1 = wages above similar workers in similar sector and region, 0 = otherwise	0.419	(0.493)
InternationalSales	1 = more than 50% of sales sold abroad, 0 = otherwise	0.241	(0.428)
Other HRM practices			
Training	1 = training provided, 0 = otherwise	0.792	(0.406)
TemporaryWorkers	Share of temporary workers	20.690	(21.220)
No Fire Measures[b]	Number of measures mentioned to avoid firing permanent workers	1.915	(0.278)
Number of observations		734	

[a]See Table A2 in the appendix for a detailed description of how variables have been constructed.
[b]For this variable, the number of observations is 178.

information available, we can distinguish among 91 different sectors. Since it is very important to analyze the provision of incentives among plants that have similar difficulties in observing effort, we include sectorial dummies in all of our regressions.

- Region

  The province in which the plants are located also appears in our dataset. There are 50 different provinces within Spain, which correspond to 17 different larger regions (autonomous communities). Although the labor legislation is exactly the same in all regions, part of the collective bargaining between unions and employers' representatives is done at a provincial/regional level (see Diaz-Moreno & Galdon-Sanchez, 2004). Therefore, it may still be important to control for possible region effects, given the existence of potential differences in the negotiation of some labor conditions between unions and employers.

- Age of the establishment

  In the dataset we also have information regarding the year in which the establishment was founded. We construct the variable *Old*, which takes value one if the establishment was founded before 1980 and zero otherwise. The year 1980 is particularly relevant in Spain since it is the year in which the Worker's Statute, the main law that regulates the different aspects of labor relations in the Spanish democratic era, was signed.

- Ownership

  Different sources of information regarding the ownership structure of the firms are available in the dataset. From this information we construct the following variables. We define the variable *StateShare*, which takes value one if the state owns a share of the firm and zero otherwise. Around three percent of firms in our sample have some of their shares owned by the state. Among these, on average, 65 percent of their capital is state owned. Moreover, since the establishments specify if they belong (totally or partially) to a multinational group, we can define the variable *Multinational*, which takes value one if the firm belongs to a multinational group and zero otherwise.

- Size

  The size of the establishment is also available since the dataset provides information regarding the number of workers employed at each establishment. We define the variable *Large* that takes value one if the firm has more than 500 workers. Otherwise, it takes value zero.

- Union

  Information on the presence and influence of unions in the firm can also be obtained from the available data. In Spain, most large firms negotiate an agreement beyond the regional pact that applies solely to that firm. All workers, unionized or not, are subject to this agreement. A unionized worker has the right to enter in this negotiation process, since unionized workers have the right to choose their representatives in the negotiation with the firm, among themselves, through voting. The number of unionized workers at the firm can play an important role in determining the type of agreement reached since this number also gives an idea of the strength of unions in the firm (see Diaz-Moreno & Galdon-Sanchez, 2004). Therefore, we specify the variable *Union* that takes value one if the level of workers' unionization is higher than 60 percent and zero in the remaining cases.
- Wage Level

  Firms are asked to compare the wages that they pay to their workers with the wages of similar workers of similar firms in the same region. We construct the variable *WageLevelAbove*, which takes value one if firms say that their workers' wages are above comparable workers' wages and zero if it is the opposite.
- Foreign Product Markets

  The dataset has information regarding the distribution of firms' sales in Spain, Europe, and the rest of the world. From this information, we designate the variable *InternationalSales* that takes value one if more than 50 percent of the firm's sales are international and zero if it is not the case.

Once we establish that seniority-based pay is used as a motivational device, and in order to provide further evidence, we analyze different factors and personnel policies that could be more relevant to the use of seniority-based pay than to the use of explicit incentives. These are described below.

- Temporary (or Fixed-Term) Contracts[14]

  The proportion of workers under fixed-term and permanent contracts is also available in the dataset. This ranges from 0 to 96 percent and the average is around 21 percent. The variable that registers the share of temporary workers is *TemporaryWorkers*.
- Firing policies

  There is information regarding firing policies from those firms that have recently fired workers or that were in a staff cutback process at the time. However, the number of observations for these variables is

reduced substantially since many firms in the sample were not undergoing a process such as this. In particular, firms were asked about the adoption of alternative policies to avoid firing workers with permanent contracts. These policies included ending temporary contracts, reducing production subcontracted to other firms, relocating multiskilled workers, cutting back or cancelling overtime, distributing labor hours (reducing hours of affected workers), and offering early retirement to older workers. Firms were asked to select the two measures that were mostly used by them. We use the information provided for the firms that were involved in this process to define the variable *NoFireMeasures*, which makes reference to the number of measures taken to avoid firing permanent workers. These could be zero (i.e, firm did not mention any measure), one (i.e, firm only mentioned one measure), or two (i.e, firm mentioned two different measures). This variable measures the degree of commitment of firms to keep a long-term relationship with their workers.

- Training

  As we have previously mentioned, wages can be correlated with worker's tenure for reasons other than incentives. A noticeable alternative to this explanation is the existence of training. We collected information on whether blue-collar workers were offered training courses, which led us to the variable *Training*. This variable takes value one if training was offered by the establishment to blue-collar workers and zero if it was not.

Table A2 in the appendix summarizes how the main variables regarding human resource management practices were defined. Table 1 provides the definition of the variables and their basic summary statistics.

4.2. Descriptive Analysis

The descriptive analysis of the variables used in our exercise can be found in Table 2. This analysis is based on the variable *Seniority-Based Pay (incidence)*. The left-hand panel of this table displays the summary statistics for the main variables used by those firms for which *Seniority-Based Pay (incidence)* is equal to one or, in other words, firms that set wages partly on seniority. The central panel displays the summary statistics for the main variables used by firms where seniority is never used as a criteria to set wages or where *Seniority-Based Pay (incidence)* is equal to zero. The right-hand

Table 2. Descriptive Statistics, by *Seniority-Based Pay (Incidence)*.

Variable	Definition	Seniority-Based Pay[a]		Non-Seniority-Based Pay [b]		*p*-Value
		Mean	SD	Mean	SD	
Explicit Incentives (incidence)	1 = explicit incentives provided, 0 = otherwise	0.578	(0.495)	0.636	(0.481)	0.069*
Explicit Incentives (intensity)	Percentage of earnings that correspond to incentive pay	9.057	(10.386)	10.627	(11.243)	0.040**
Explicit Individual Incentives (incidence)	1 = only explicit individual incentives provided, 0 = otherwise	0.493	(0.501)	0.530	(0.499)	0.183
Explicit Individual Incentives (intensity)	Percentage of earnings that correspond to individual incentive pay	8.019	(10.167)	9.618	(11.470)	0.039**
Monitoring (incidence)	1 = workers subject to high supervision, 0 = otherwise	0.364	(0.482)	0.409	(0.492)	0.134
Monitoring (intensity)	Level of supervision (1 to 5)	3.331	(0.699)	3.377	(0.701)	0.071*
Old	1 = plant founded before 1980, 0 = otherwise	0.819	(0.385)	0.705	(0.456)	0.000***
StateShare	1 = state owns a share of the firm, 0 = otherwise	0.061	(0.241)	0.022	(0.149)	0.004**
Multinational	1 = firm belongs to multinational group, 0 = otherwise	0.289	(0.454)	0.284	(0.451)	0.454
Large	1 = firm with more than 500 workers, 0 = otherwise	0.156	(0.364)	0.087	(0.283)	0.003**
Union	1 = unionization of workers above 60%, 0 = otherwise	0.327	(0.470)	0.315	(0.465)	0.381
WageLevelAbove	1 = wages above similar workers in similar sector and region, 0 = otherwise	0.426	(0.495)	0.416	(0.493)	0.404
InternationalSales	1 = more than 50% of sales sold abroad, 0 = otherwise	0.218	(0.413)	0.250	(0.433)	0.176
Training	1 = training provided, 0 = otherwise	0.767	(0.423)	0.801	(0.399)	0.157
TemporaryWorkers	Share of temporary workers	16.590	(19.100)	22.350	(21.816)	0.000***
No Fire Measures	Number of measures mentioned to avoid firing permanent workers	1.942	(0.235)	1.899	(0.303)	0.158
Number of observations		211		523		

Note: ***p-value < 0.01, **p-value < 0.05, *p-value < 0.1.
[a]Seniority-Based Pay (incidence variable) = 1.
[b]Seniority-Based Pay (incidence variable) = 0.

panel displays the *p*-values associated with the one-sided tests regarding the difference in variable means for firms that base wages partly on seniority and those that do not.

The first important feature to notice is that the firms that base wages partly on seniority are less likely to provide explicit incentives than those that do not provide such wage scheme. This is true for both the incidence and the intensity measures of explicit incentives. These firms also tend to undertake less monitoring in terms of both our measures (incidence and intensity), although the difference is not significant in the case of the incidence measure. These factors provide some preliminary evidence that seniority-based pay and other incentive mechanisms can be considered substitutive devices. It is also worth noting that these firms tend to be older, partly or totally owned by the state, and larger. Firms that offer wages according to seniority tend to be more unionized, although the difference is not significant. Since the firm's characteristics could affect the way in which the firm sets its wages, it is important to control for these factors in our regression analysis. For example, state-owned and/or large firms may have a preference for rules rather than discretion with regard to their pay schemes. Therefore, it is important to see if the negative relationship between seniority-based pay and explicit incentives, which appears in the raw data, stays the same once these variables are included as controls in our analysis.

With regard to other personnel policies, firms that base wages partly on seniority have also a lower proportion of workers under fixed-term contracts. Also, on average, they reported a higher number of measures taken to avoid firing their core permanent workers, although the difference is not significant. Regarding training and seniority-based pay, Table 2 shows that there is no difference between firms that base wages partly on seniority and those that do not in terms of training.

5. RESULTS

In this section we undertake the empirical analysis and explain the results obtained. We want to explore if a firm's predisposition to offer wages partly based on seniority is related to long-term incentives; that is, to find out if the negative correlation between seniority-based pay and other incentive devices (explicit incentives or monitoring devices) remains after controlling for different firm characteristics as well as regional and sectorial controls.[15]

In particular, we estimate a probit model in which *Seniority-Based Pay (incidence)* is the dependent variable, and include the incidence measures of

explicit incentives and monitoring as regressors. Then we estimate an ordered probit model in which *Seniority-Based Pay (intensity)* is the dependent variable and include the intensity measures of explicit incentives and monitoring.[16] The results are displayed in Tables 3a and 3b, respectively.

We start with the most simple specification that includes *Explicit Incentives (incidence)* as an explanatory variable as well as the mentioned controls. As column (1) of Table 3a indicates, firms that base wages partly on seniority are less likely to offer explicit incentives, even after controlling for different firm characteristics. This result confirms the findings of Barth (1997). Working with a sample of Norwegian workers, he found that piece-rate workers have a negligible return to seniority in terms of wages. Column (2) analyzes the relationship between seniority-based pay and monitoring. Again, a negative relationship remains after controlling for firm characteristics. This result is similar to the findings of Hutchens (1987). Using US

Table 3a. Seniority-Based Pay, Explicit Incentives, and Monitoring Incidence Measures: Probit Estimates.

Dependent Variable: Seniority-Based Pay (Incidence)	(1)	(2)	(3)	(4)	(5)
Explicit Incentives (incidence)	−0.296**		−0.412**		
	(0.127)		(0.162)		
Explicit Individual Incentives (incidence)				−0.199	−0.468***
				(0.123)	(0.160)
Monitoring (incidence)		−0.269**	−0.465**		−0.640***
		(0.124)	(0.208)		(0.187)
Explicit Incentives × Monitoring			0.318		
			(0.259)		
Explicit Individual Incentives × Monitoring					0.689**
					(0.252)
CONTROLS	Yes	Yes	Yes	Yes	Yes
SECTOR DUMMIES	Yes	Yes	Yes	Yes	Yes
REGION DUMMIES	Yes	Yes	Yes	Yes	Yes
Log likelihood	−347.895	−348.249	−344.934	−349.315	−343.26
χ^2	112.8	112.1	118.73	109.97	122.07
Number of observations	654	654	654	654	654

Notes: Columns (4) and (5) are equivalent to columns (1) and (2) using the measure *Explicit Individual Incentives*.

Controls include *Old, StateShare, Multinational, Large, Union, WageLevelAbove, InternationalSales*.

Standard errors in parenthesis, ***p-value < 0.01, **p-value < 0.05, *p-value < 0.1.

Table 3b. Seniority-Based Pay, Explicit Incentives and Monitoring Intensity Measures: Ordered Probit Estimates.

Dependent Variable: Seniority-Based Pay (Intensity)	(1)	(2)	(3)	(4)	(5)
Explicit Incentives (intensity)	−0.013**		−0.013**		
	(0.005)		(0.005)		
Explicit Individual Incentives (intensity)				−0.010*	−0.010*
				(0.005)	(0.005)
Monitoring (intensity)		−0.141*	−0.141*		−0.137*
		(0.081)	(0.082)		(0.081)
CONTROLS	Yes	Yes	Yes	Yes	Yes
SECTOR DUMMIES	Yes	Yes	Yes	Yes	Yes
REGION DUMMIES	Yes	Yes	Yes	Yes	Yes
Ancillary parameter 1	0.671	0.318	0.160	0.678	0.139
	(0.494)	(0.574)	(0.579)	(0.494)	(0.580)
Ancillary parameter 2	1.634	1.278	1.126	1.640	1.106
	(0.496)	(0.575)	(0.580)	(0.496)	(0.581)
Log likelihood	−487.036	−488.235	−485.554	−487.895	−486.466
χ^2	162.08	159.68	165.04	160.36	163.22
Number of observations	734	734	734	734	734

Notes: Columns (4) and (5) are equivalent to columns (1) and (2) using the measure *Explicit Individual Incentives*.
Controls include *Old, StateShare, Multinational, Large, Union, WageLevelAbove, InternationalSales*.
Standard errors in parenthesis, ***p-value < 0.01, **p-value < 0.05, *p-value < 0.1.

data, he proved that monitoring difficulties correlate positively with the application of deferred payment schemes.

Jobs that offer piece-rate payments are subject to indirect monitoring (see Lazear, 1979). As Hutchens (1987) clearly explains, in this case, monitoring essentially takes the form of counting the units produced, so workers are paid accordingly. In column (3), we apply both *Explicit Incentives (incidence)* and *Monitoring (incidence)* as right-hand side variables. Moreover, we allow an interaction term between these two variables. The coefficient on these two variables remains negative in this specification. The coefficient of the interaction term is not statistically different from zero, suggesting that there is no additional effect coming from firms that invest in monitoring devices at the same time they provide incentives.

Table 3b reports the results when repeating the previous exercise but using intensity measures instead. According to the overall analysis, results are qualitatively the same.[17] All these results suggest that seniority-based pay

and explicit incentives, as well as monitoring devices, act as substitutes. This suggests that seniority-based pay is used as a motivational device, in accordance with the main prediction of Lazear's theory. The intuition is simple: the more difficult a job is to supervise and the less resources devoted by the firm to control its workers, the more likely the firm relies on seniority-based pay.

As mentioned earlier, different personnel practices are chosen simultaneously by a firm. One possible way of solving this simultaneity problem is to estimate multivariate probits of the different incentive practices. In this case, the correlation coefficient between the different equations captures the relationship between the different practices. We estimate bivariate probit models in which *Seniority-Based Pay (incidence)* and *Explicit Incentives (incidence)*, and *Seniority-Based Pay (incidence)* and *Monitoring (incidence)*, respectively, are the dependent variables (columns (1) and (2) in Table 4). We also estimate a trivariate probit model in which *Seniority-Based Pay (incidence)*, *Explicit Incentives (incidence)*, and *Monitoring (incidence)* are the dependent variables (column (3) in Table 4).[18] As it can be appreciated, the correlation coefficients between the variables *Seniority-Based Pay (incidence)* and *Explicit Incentives (incidence)*, and *Seniority-Based Pay (incidence)* and *Monitoring (incidence)* are negative and significant, providing further evidence that these are substitutive practices. The correlation coefficient between the variables *Explicit Incentives (incidence)* and *Monitoring (incidence)* is positive but not significant.

Once we established that seniority-based pay is a substitute for other motivational devices, we further analyze the relationship between this policy and other personnel practices. Economic theory suggests that firms that decide to use seniority-based pay, rather than explicit incentives, as an incentive device should complement such policy with other personnel practices that give the firm the necessary credibility to commit to future wages. As Hutchens (1987) states, seniority-based pay contracts should be accompanied by long job tenure. However, this should not be the case for firms that offer explicit incentives. Next, we study the combination of different personnel practices. We start analyzing the use of short-duration contracts or temporary contracts. Table 5a displays the estimates of a trivariate probit model in which *Seniority-Based Pay (incidence)*, *Explicit Incentives (incidence)* and *Monitoring (incidence)* are the dependent variables. We include the share of temporary contracts in the firm, *TemporaryWorkers*, as independent variable.

Overall, the results in Table 5a show that firms that opt for seniority-based pay are less likely to use short duration contracts.[19] This result

Table 4. Seniority-Based Pay, Explicit Incentives, and Monitoring Incidence Measures: Bivariate and Trivariate Probit Estimates.

Dependent Variables	Seniority-Based Pay and Explicit Incentives	Seniority-Based Pay and Monitoring	Seniority-Based Pay, Explicit Incentives, and Monitoring		Seniority-Based Pay and Explicit Individual Incentives	Seniority-Based Pay, Explicit Individual Incentives, and Monitoring	
	(1)[a]	(2)[a]	(3)[b]		(4)[a]	(5)[b]	
	Seniority-Based Pay	*Seniority-Based Pay*	*Seniority-Based Pay*	*Monitoring*		*Seniority-Based Pay*	*Monitoring*
Correlation Coefficients							
Explicit Incentives	−0.197**		−0.193**	0.089			
	(0.074)		(0.074)	(0.073)			
Explicit Individual Incentives					−0.128**	−0.124*	0.039
					(0.075)	(0.073)	(0.071)
Monitoring		−0.168**	−0.158**			−0.166**	
		(0.073)	(0.074)			(0.076)	
CONTROLS[c]	Yes	Yes		Yes	Yes		Yes
SECTOR DUMMIES	Yes	Yes		Yes	Yes		Yes
REGION DUMMIES	Yes	Yes		Yes	Yes		Yes
Log likelihood	−729.913	−759.300	−1138.631		−754.651	−1163.374	
χ^2	196.03	190.25	291.32		219.17	315.37	
Number of observations	734	734	734		734	734	

Notes: Columns (4) and (5) are equivalent to columns (1) and (2) using the measure *Explicit Individual Incentives*.
Standard errors in parenthesis, ***p-value<0.01, **p-value<0.05, *p-value<0.1.
[a]Bivariate Probit.
[b]Trivariate Probit. Simulated maximum-likelihood estimates using GHK smooth recursive simulator (100 random draws).
[c]As in Table 3a.

Table 5a. Seniority-Based Pay, Explicit Incentives, Monitoring, and Temporary Contracts Incidence Measures: Trivariate Probit Estimates[a].

	Panel A			Panel B		
Dependent variables	Seniority-Based Pay, Explicit Incentives, and Monitoring[b]			Seniority-Based Pay, Explicit Individual Incentives, and Monitoring[b]		
	Seniority-Based Pay	*Explicit Incentives*	*Monitoring*	*Seniority-Based Pay*	*Explicit Individual Incentives*	*Monitoring*
Estimation coefficients						
TemporaryWorkers	−0.009*	0.003	0.005*	−0.009*	0.003	0.005*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
CONTROLS[c]		Yes			Yes	
SECTOR DUMMIES		Yes			Yes	
REGION DUMMIES		Yes			Yes	
	Seniority-Based Pay	*Monitoring*		*Seniority-Based Pay*	*Monitoring*	
Correlation coefficients						
Explicit Incentives	−0.192**	0.086				
	(0.075)	(0.073)				
Explicit Individual Incentives				−0.138*	0.042	
				(0.074)	(0.072)	
Monitoring	−0.146**			−0.155**		
	(0.075)			(0.075)		
Log likelihood		−1138.957			−1157.958	
χ^2		297.34			322.09	
Number of observations		734			734	

Notes: Panel B is equivalent to Panel A using the measure *Explicit Individual Incentives*.
Standard errors in parenthesis, ***p-value<0.01, **p-value<0.05, *p-value<0.1.
[a]Simulated maximum-likelihood estimates using GHK smooth recursive simulator (100 random draws).
[b]Incidence Measures.
[c]As in Table 3a.

indicates a commitment to long employment relations that can be explained in terms of the incentive role of seniority-based pay practices. On the other hand, as expected, the coefficient of the variable *TemporaryWorkers* is not generally significant in the equations in which the dependent variables are *Explicit Incentives (incidence)* or *Monitoring (incidence)*. The correlation coefficients between these variables are similar, although smaller, to those reported in Table 4.

We next analyze different firing policies. As mentioned earlier, these variables have fewer observations because only the firms that had been recently under a process of restructuring had to answer the questions related to firing policies. For this reason, we are forced to estimate separate probit models (rather than a trivariate probit) in which *Seniority-Based Pay (incidence)*, *Explicit Incentives (incidence)*, and *Monitoring (incidence)* are the dependent variables.[20] Table 5b displays these estimates.[21]

The first column of Table 5b shows that firms that base wages partly on seniority are more likely to implement measures to avoid firing their permanent workers. The second column reveals that for firms that use *Explicit Incentives (incidence)*, the coefficient on *NoFireMeasures* is positive but not significant. Finally, the relationship between *Monitoring (incidence)* and *NoFireMeasures* is negative, although not significant. Overall these

Table 5b. Seniority-Based Pay, Explicit Incentives, Monitoring and No Firing Measures Incidence Measures: Probit Estimates.

Dependent Variable	Seniority-Based Pay (1)	Explicit Incentives (2)	Monitoring (3)	Explicit Individual Incentives (4)
No Fire Measures	2.611**	0.785	−0.377	−0.173
	(1.191)	(0.809)	(0.627)	(0.682)
CONTROLS[a]	Yes	Yes	Yes	Yes
SECTOR DUMMIES	Yes	Yes	Yes	Yes
REGION DUMMIES	Yes	Yes	Yes	Yes
Log likelihood	−42.743	−42.376	−66.137	−46.399
χ^2	65.12	50.74	30.03	42.89
Number of observations	111	98	120	62

Notes: Column (4) is equivalent to column (2) using the measure *Explicit Individual Incentives*. Standard errors in parenthesis, ***p-value < 0.01, **p-value < 0.05, *p-value < 0.1.
[a]As in Table 3a.

results suggest that firms that choose to base wages partly on seniority also choose other personnel practices that involve long employment relationships, which is consistent with the idea that seniority-based pay is used to provide long-term incentives.

As mentioned earlier, there are alternative theories that predict a positive relationship between wages and seniority for reasons other than the provision of incentives. In particular, this could be the case in the presence of training policies. We estimate a probit model in which the dependent variable is *Seniority-Based Pay (incidence)* and an ordered probit in which the dependent variable is *Seniority-Based Pay (intensity)*, as follows. In this model, we include the variable *Training* as an explanatory variable. Table 6 displays the results of this exercise.

The main result in this exercise is that training and seniority-based pay are negatively related. This suggests that firms that base wages partly on seniority are not more likely to train their workers than firms that do not pay according to accumulated tenure. Several clarifications are worth noting. First, the variable *Training* is a general measure of training and not necessarily training on firm-specific skills. Second, this variable captures training activities from the previous year and not overall training activities or training required in the current job.[22]

Of course, our findings do not rule out training as a mechanism that generates a positive correlation between wages and seniority or the fact that trained workers may receive higher wages due to their tenure. Instead, our sample suggests – keeping these clarifications in mind – that there are

Table 6. Seniority-Based Pay and Training.

Dependent Variable	Seniority-Based Pay (incidence) Probit (1)	Seniority-Based Pay (intensity) Ordered Probit (2)
Training	−0.262*	−0.304**
	(0.157)	(0.146)
CONTROLS[a]	Yes	Yes
SECTOR DUMMIES	Yes	Yes
REGION DUMMIES	Yes	Yes
Log likelihood	−349.221	−487.584
χ^2	110.15	160.98
Number of observations	654	663

Standard errors in parenthesis, ***p-value < 0.01, **p-value < 0.05, *p-value < 0.1.
[a]As in Table 3a.

reasons beyond training that explain the practice of seniority-based pay. In this chapter, we have argued that there is evidence that seniority-based pay is used as an incentive device.

6. CONCLUSIONS

In this chapter, we have empirically tested the theory of long-term implicit contracts using plant-level data. In particular, we have analyzed the possible motivation role of seniority-based pay schemes. Unlike previous papers, we have used a direct measure of such firm practice.

Our main conclusion is that firms that base wages partly on seniority are less likely to offer explicit incentives. They are also less likely to invest in monitoring devices. This result remains stable after controlling for several firm characteristics. Another interesting result that arises from our exercise is that firms that base wages partly on seniority are more likely to engage in other personnel practices that involve long employment relationships. These practices make the firm's commitment to pay high future wages credible and therefore are complementary measures to implicit incentives. Overall, our plant-level data provide empirical support to the implicit incentives theory proposed by Lazear (1979).

On the other hand, we think that, in order to properly test personnel economics theories, plant-level data on the firm's practices are required. Even though the data are costly to gather and have so far been scarce, they contain valuable information that can shed new light on testing personnel economic theories.

NOTES

1. Section 2 reviews these theories in detail.
2. A more recent test of explicit incentives is provided in Lazear (2000).
3. The next section includes a review of the literature.
4. Bayo-Moriones and Huerta-Arribas (2002a, 2002b) have studied explicit incentives using the same dataset that we use here. In Bayo-Moriones and Huerta-Arribas (2002a), the authors investigate the factors that influence the adoption of incentive schemes that link the blue-collar workers pay to the results achieved by the establishment that employ them, i.e. the so-called organizational incentive plans. And in Bayo-Moriones and Huerta-Arribas (2002b), they identify the factors that determine the use of production incentives for manual workers in the Spanish manufacturing industry.

5. Notice that for such threat to be credible, some form of monitoring, which allows firms to obtain at least a qualitative measure of worker's performance, has to be feasible. While a necessary condition for firms to implement explicit incentives is that output is easy to observe and thus monitoring allows to quantify output and pay accordingly, a more imperfect form of monitoring is sufficient to implement seniority-based pay.

6. Note that this explanation does not rely on the existence of a seniority-based pay contract. However, it relies on the presence of training at the firm, something we are able to analyze using our data.

7. This is equivalent to the ISIC rev. 3 activity classification.

8. As Osterman (1994) states: "The great advantage of surveying establishments, as opposed to firms, is that the respondent in an establishment is likely to know the facts" (page 174).

9. This refers to all types of workers in the firm.

10. These correspond to 75 percent of the original sample and thus sizeable selection problems should not be at work.

11. Employment protection legislation in Spain is very similar to most European countries. Firms can fire workers for "economic reasons" (in which case the worker gets an indemnity) or for "disciplinary reasons" (in which case the worker has no right to an indemnity). Workers can always appeal the case if they disagree. If a dismissal case ends up in court, firms may have to pay larger indemnities to workers. Therefore, while dismissal is costly for the firm, it is still a possibility even for permanent workers.

12. As explained in Section 2.

13. We have not been able to find any other paper that studies this variable with a cross section of firms, so we cannot establish any comparison.

14. In 1984, there was a reform of the Spanish Labor Law that allowed the use of fixed-term contracts for jobs whose nature was not necessarily temporary. These contracts involve much lower termination costs than permanent contracts (see, for instance, Güell, 2000; Alonso-Borrego, Fernandez-Villaverde, & Galdon-Sanchez, 2005, for an analysis of their effect in the Spanish economy).

15. The coefficients of the different controls are available upon request.

16. Around 10 percent of the observations in the sample show no variation in terms of the dependent variable within sectors or regions and are lost when estimating the probit model.

17. The estimation of an ordered probit of seniority-based pay (intensity) on incidence measures of explicit incentives and monitoring leads to the same findings. These results are available on request.

18. For estimation of this type of models see, for instance, Cappelari and Jenkins (2003).

19. Using incidence measures of temporary contracts, we obtain the same qualitative results. Also, ordered probit estimates in which *Seniority-Based Pay (intensity)* is the dependent variable provide similar results. These results are available upon request.

20. In these estimations, in order to maximize the number of observations, regional dummies correspond to the 17 autonomous communities instead of the 50 provinces.

21. Between 32 and 44 percent of the observations in the sample show no variation in terms of the dependent variable within sectors or regions and are lost when estimating the probit model.

22. Barth (1997) has information on the job's required level of on-the-job training. He finds that firm-specific training has a negative effect on the tenure wage profile.

ACKNOWLEDGMENTS

The authors are grateful to Ghazala Azmat, Erling Barth, Mike Gibbs, Paolo Ghinetti, Stepan Jurajda, Marco Manacorda, Pedro Ortin, as well as seminar participants at Universidad Carlos III de Madrid, the Workshop on the Use and Analysis of Employer-Employee Data at the Institute for Social Research (Oslo), European Economic Association Annual Congress, the Conference on the Analysis of Firms and Employees, and an anonymous referee for very useful comments and suggestions. The authors would like to express their gratitude to Fundacion BBVA for providing the means to create the database used in this study. Bayo-Moriones, Galdon-Sanchez and Güell acknowledge financial support from Ministerio de Educacion y Ciencia, projects SEJ2007-66511, ECO2008-02641 and SEJ2006-09993/ECON, respectively. Güell acknowledges the support of the Barcelona GSE Research Network and of the Government of Catalonia. Galdon-Sanchez and Güell thank the hospitality of the CEP at LSE where part of this work was completed.

REFERENCES

Alonso-Borrego, C., Fernandez-Villaverde, J., & Galdon-Sanchez, J. E. (2005). *Evaluating labor market reforms: A general equilibrium approach*. NBER Working Paper 11519. National Bureau of Economic Research, Cambridge, MA, USA.

Baron, J. N., & Kreps, D. M. (1999). *Strategic human resources*. New York: Wiley.

Barth, E. (1997). Firm-specific seniority and wages. *Journal of Labor Economics*, *15*(3), 495–506, Pt. 1.

Bayo-Moriones, A., & Huerta-Arribas, E. (2002a). Organisational incentive plans in Spanish manufacturing industry. *Personnel Review*, *31*(2), 128–142.

Bayo-Moriones, A., & Huerta-Arribas, E. (2002b). The adoption of production incentives in Spain. *British Journal of Industrial Relations*, *40*(4), 709–724.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis with special reference to education*. New York: National Bureau of Economic Research.

Brown, J. N. (1989). Why do wages increase with tenure? On-the-job training and life-cycle wage growth observed within firms. *American Economic Review*, *79*(5), 478–498.

Cappelari, L., & Jenkins, S. (2003). *Multivariate probit regression using simulated maximum likelihood*. Mimeographed document ISER. University of Essex.

Clark, R. L., & Ogawa, N. (1992). The effect of mandatory retirement on earnings profiles in Japan. *Industrial and Labor Relations Review*, *45*(2), 258–266.

Diaz-Moreno, C., & Galdon-Sanchez, J. E. (2004). Collective bargaining under complete information. In: S. W. Polachek (Ed.), *Accounting for worker well-being, research in labor economics* (Vol. 23, pp. 359–379). Elsevier Science/JAI Press.

Felli, L., & Harris, C. (1996). Learning, wage dynamics, and firm-specific human capital. *Journal of Political Economy*, *104*(4), 838–868.

Flabbi, L., & Ichino, A. (2001). Productivity, seniority and wages: New evidence from personnel data. *Labour Economics*, *8*(3), 359–387.

Frank, R. H., & Hutchens, R. M. (1993). Wages, seniority, and the demand for rising consumption profiles. *Journal of Economic Behavior and Organization*, *21*(3), 251–276.

Freeman, R., & Medoff, J. (1984). *What do unions do?* New York: Basic Books.

Galdon-Sanchez, J. E., & Güell, M. (2003). Dismissal contracts and unemployment. *European Economic Review*, *47*(2), 127–139.

Güell, M. (2000). *Fixed-term contracts and unemployment: An efficiency wage analysis*. Working Paper no. 433. Industrial Relations Section, Princeton University, Princeton, NJ, USA.

Harris, M., & Holmstrom, B. (1982). Ability, performance and wage differentials. *Review of Economic Studies*, *49*(3), 315–333.

Hellerstein, J. K., & Neumark, D. (1995). Are earnings profiles steeper than productivity profiles? Evidence from Israeli firm-level data. *Journal of Human Resources*, *30*(1), 89–112.

Holmstrom, B., & Milgrom, P. (1994). The firm as an incentive system. *American Economic Review*, *84*(4), 972–991.

Hutchens, R. M. (1987). A test of Lazear's theory of delayed payment contracts. *Journal of Labor Economics*, *5*(4), S153–S170, Pt. 2.

Hutchens, R. M. (1989). Seniority, wages and productivity: A turbulent decade. *Journal of Economic Perspectives*, *3*(4), 49–64.

Ichniowski, C., & Shaw, K. (2003). Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices. *Journal of Economic Perspectives*, *17*(1), 155–178.

Idson, T. L., & Valletta, R. G. (1996). Seniority, sectoral decline, and employee retention: An analysis of layoff unemployment spells. *Journal of Labor Economics*, *14*(4), 654–676.

Kotlikoff, L. J., & Gohkale, J. (1992). Estimating a firm's age-productivity profile using the present value of workers' earnings. *Quarterly Journal of Economics*, *107*(4), 1215–1242.

Lazear, E. P. (1979). Why is there mandatory retirement? *Journal of Political Economy*, *87*(6), 1261–1264.

Lazear, E. P. (1981). Agency, earnings profiles, productivity, and hours restrictions. *American Economic Review*, *71*(4), 606–620.

Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review*, *90*(5), 1346–1361.

Lazear, E. P., & Moore, R. L. (1984). Incentives, productivity, and labor contracts. *Quarterly Journal of Economics*, *99*(2), 275–296.

Levine, D. I. (1993). Worth waiting for? Delayed compensation, training, and turnover in the United States and Japan. *Journal of Labor Economics*, *11*(4), 724–752.

Loewenstein, G., & Sicherman, N. (1991). Do workers prefer increasing wage profiles? *Journal of Labor Economics*, *9*(1), 67–84.

Medoff, J. L., & Abraham, K. G. (1980). Experience, performance, and earnings. *Quarterly Journal of Economics*, *95*(4), 703–736.
Mincer, J. (1974). *Schooling, experience, and earnings*. New York: National Bureau of Economic Research.
Osterman, P. (1994). How common is workplace transformation and who adopts it? *Industrial and Labor Relations Review*, *47*(2), 173–188.
Salop, J., & Salop, S. C. (1976). Self-selection and turnover in the labor market. *Quarterly Journal of Economics*, *90*(4), 619–627.
Spitz, J. (1991). *Productivity and wage relations in economic theory and labor markets*. Ph.D. Dissertation, Stanford University Graduate School of Business.

APPENDIX

Table A1. Ratio Sample to Population, by Firm Size and Sector.

Sector/Firm Size	50–199	200–499	500 or more	Total
Food, drinks, and tobacco	12.53	16.67	28.85	14.09
Textiles, clothing, leather goods, and footwear	13.94	19.48	54.55	15.05
Wood and cork	17.36	18.18	0.00	17.31
Paper, publishing, and graphic arts	12.65	20.29	55.56	14.52
Chemical industry	14.75	10.17	22.86	14.23
Rubber and plastics	14.56	24.24	28.57	15.98
Nonmetallic mineral products	13.27	11.76	40.00	13.61
Primary metal industries and fabricated metal products	15.88	13.13	48.28	16.83
Machinery and mechanical equipment	13.83	19.05	42.11	15.72
Electrical material and equipment, electronics, and optics	13.67	19.59	39.39	17.16
Transport material	17.70	26.44	48.21	24.39
Miscellaneous manufacturing industries	20.33	28.57	66.67	21.69
Total	14.44	17.63	38.97	16.05

Note: Sector corresponds to the Spanish equivalent to ISIC (CNAE).

Table A2. HRM Practices: Variables Description.

Survey Questions (Q) and Answers (A)	Variable Name	Variable Values
Q1: On which of these factors does the fixed part of the wage of manual workers at this plant most closely depend? Which comes in second? **A:** Type of job, skill level, seniority, efficiency of their work, personal assessment from supervisor.	*Seniority-Based Pay (incidence)*	1 = seniority mentioned either as the most important or the second most important factor when setting wages, 0 = otherwise.
	Seniority-Based Pay (intensity)	2 = seniority mentioned first, 1 = seniority mentioned second, 0 = otherwise.
Q2: Do the manual workers at this plant receive any type of incentive payment? **A:** Yes, No.	*Explicit Incentives (incidence)*	1 = Yes, 0 = otherwise.
Q3: What type of incentives? **A:** Based on productivity; on quality; on plant-level or firm's results; other types.	*Explicit Individual Incentives (incidence)*	1 = Yes on Q2 and based on productivity and/or quality on Q3, 0 = otherwise (this includes firms answering 'No' in Q2; also firms using both individual incentives (based on productivity and/or quality) as well as collective incentives (based on plant-level or firm's results and/or other types).
Q4: Among those manual workers who receive incentives, what percentage of their earnings (on average) represents such incentives? **A:** %	*Explicit Incentives (intensity)*	% number.
	Explicit Individual Incentives (intensity)	% number for firms which *Explicit Individual Incentives (incidence)* equals 1.
Q5: Which of the following phrases best describes the degree of supervision to which your employees are subject? **A:** 1. No supervision at all; 2. Hardly any supervision; 3. Moderate supervision; 4. Quite close supervision; 5. Close supervision.	*Monitoring (incidence)*	1 = high and very high supervision, 0 = otherwise.
	Monitoring (intensity)	(1, 5)
Q6: On average, how many hours of training per worker were given last year? **A:** Number of hours.	*Training*	1 = number of hours is positive, 0 = otherwise.
Q7: How many permanent and temporary workers were employed at your plant at the end of last year? **A:** Number of workers.	*TemporaryWorkers*	Share of temporary workers (0, 100).

Table A2. (*Continued*)

Survey Questions (Q) and Answers (A)	Variable Name	Variable Values
Q8: When downsizing is in progress, measures are usually taken to avoid laying off permanent workers. Among the measures of this type that appear on this card, which are the main ones that have been adopted/are planned to be adopted in the downsizing of the workforce at this plant? (Please, select the two main ones and rank them in order of importance.) **A:** 1. Ending temporary contracts; 2. Reducing production subcontracted to other firms; 3. Relocating multiskilled workers; 4. Cutting back or cancelling overtime; 5. Distributing labor hours (reducing hours of affected workers); 6. Offering early retirement to older workers.	*NoFireMeasures*	Number of measures mentioned to avoid firing permanent workers (0,2).

THE PROMOTION DYNAMICS OF AMERICAN EXECUTIVES

Christian Belzil and Michael Bognanno

ABSTRACT

We formulate static and dynamic empirical models of promotion where the current promotion probability depends on the hierarchical level in the firm, individual human capital, unobserved individual specific attributes, time-varying firm-specific variables, as well as endogenous past promotion histories (in the dynamic version). Within the static versions, we investigate the relative influence of the key determinants of promotions and how these influences vary by hierarchical levels. In the dynamic version of the model, we examine the causal effect of past speed of promotion on promotion outcomes. The model is fit on an eight-year panel of 30,000 American executives employed in more than 300 different firms. The stochastic process generating promotions may be viewed as a series of promotion probabilities which become smaller as an individual moves up in the hierarchy and which are primarily explained by unobserved heterogeneity and promotion opportunities. Firm variables and observed human capital variables (age, tenure, and education) play a surprisingly small role. We also find that, conditional on unobservables, the promotion probability is only enhanced by the speed of promotion achieved in the past (a structural fast track effect) for a subset of the population and is negative for the majority. In general, the magnitude of the

Jobs, Training and Worker Well-Being
Research in Labor Economics, Volume 30, 189–231
Copyright © 2010 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 0147-9121/doi:10.1108/S0147-9121(2010)0000030009

individual-specific effect of past speed of promotion is inversely related to schooling, tenure, and hierarchical level.

1. INTRODUCTION

This chapter estimates static and dynamic reduced-form models of promotion using a multi-firm panel of senior executives employed at more than 300 large US corporations between 1981 and 1988. The chapter's central focus is to develop an empirical model of the determinants of promotion probability (defined as the probability of a movement in reporting level toward the rank of CEO) given: (a) the executive's human capital (age, tenure, education), (b) the firm's scale variables (profits, sales, and size), (c) the executive's firm- and level-specific promotion opportunities (the promotion rate in the firm of executives one level above the given executive), (d) the executive's reporting level in the firm's hierarchy, (e) unobserved heterogeneity (unmeasured individual and firm characteristics), and (f) the effect of the speed of the worker's past hierarchical advancement on the prospects for current advancement (estimated in a dynamic model and deemed a "fast track" when the speed of past advancement increases the probability of current advancement).

Our static model shows that the most influential factors explaining the probability of a promotion are (in order of importance): unobserved heterogeneity, the executive's reporting level in the firm, and the executive's promotion opportunities. Firm scale variables (profit, sales, and size) are of lesser importance. Finally, the executive's observed human capital variables (age, tenure, and education) play a slight role in explaining promotion probability.

Considering the first of the three main factors, unobserved heterogeneity captures the influence on promotion of persistent, difficult-to-measure individual attributes, such as ability or learning aptitude, as well as unmeasured firm characteristics that produce persistent firm-level differences in the rates of promotion. We find that the importance of unobserved heterogeneity grows with advancement in level.

The second most influential factor, reporting level, establishes that the promotion probability is not independent of constraints imposed by the executive's level in the firm's hierarchy. In both the raw data and the static model, it is clear that promotion probabilities become smaller at more senior levels. Next, the promotion opportunity variable measures the promotion

rate of executives in the level above the given executive to proxy firm- and level-specific promotion opportunities. To the extent that firm hierarchies are rigid, individual promotion outcomes will not be independent of the existence of vacancies in the hierarchy. We find that promotion opportunities play an important role in promotion, though more so at lower levels. The combined importance of level and promotion opportunities indicates that the promotion process is driven to a material extent by factors beyond the individual's control or ability, given their position in the hierarchy. Even under the assumption that all unobserved heterogeneity is individual specific, less than half of the promotion process is determined by individual factors.

Firm scale variables (profits, sales, and size (employment)) are all positively related to promotion probabilities, but they play a lesser role. Within the grouping of human capital variables, it is interesting to note that promotion probabilities are increasing in education and decreasing in age, while tenure is statistically insignificant. However, these human capital variables as a group are of little importance in predicting promotion outcomes.

In the second contribution of the chapter, a dynamic version of the model examines the effects of the past speed of promotion on current promotion probabilities after conditioning on unobserved heterogeneity. After conditioning on a worker's innate ability (unobservable heterogeneity) in the econometric model, the speed of past advancement in level negatively influences subsequent advancement for most executives. For a minority of executives, past speed of advancement aids promotion (and a fast track is found) and is associated with executives at lower levels and with lesser human capital (less education and less tenure). We believe that this finding is consistent with the hypothesis that the signaling aspect of past promotions is stronger for those who are less educated and who are relatively new in a firm. This is consistent with a job assignment model incorporating asymmetric learning. The overall influence of the speed of past promotion on subsequent promotion is negligible.

A similar finding is evident in a simple examination of the raw data in that slight evidence of a fast track operating at less senior levels in the firm is found. In comparison to others in their firm and level, younger than average executives have a slightly greater incidence of promotion at reporting levels four through six (the CEO is reporting level one). In other words, younger executives, having climbed to their level faster, are advantaged in subsequent promotions in level supporting the notion of a fast track. Fast tracks are not evident at more senior reporting levels.

A theoretical motivation for examining fast tracks in a dynamic model comes from the job assignment literature (see the Theory Appendix for more detail). Within this literature, we highlight implications drawn from two classes of models: the case of full information (e.g., Gibbons & Waldman, 1999) and the case of asymmetric learning (e.g., Waldman, 1984a; Bernhardt, 1995). In models with asymmetric learning of worker ability, the current employer is fully informed and outside firms learn worker ability through the signal provided by observing the workers job assignment. Job assignment models with heterogeneous workers, assuming either full information or asymmetric learning, imply serial correlation in promotion outcomes (fast tracks) due to differences in worker ability.

However, in models with asymmetric learning, past promotions also have an inherent effect on promotion outcomes after conditioning on worker heterogeneity. Higher wages must be paid to workers whose promotions signal high ability to outside firms. Since workers who have been rapidly promoted in the past have already been signaled to be of high ability, their subsequent promotion is less costly and, hence, speedy past promotions will have a positive causal effect on the probability of subsequent promotion. This implies that serial persistence in individual promotion histories may simultaneously result from both persistent unobserved heterogeneity and state dependence explained by past promotion outcomes.[1] An empirical analysis of fast tracks that accounts for both the dynamics of the stochastic process generating promotions and the endogeneity of the initial conditions may provide insight into both the source of fast tracks and the existence of job assignment signaling.

1.1. Promotion in the Literature

Promotion within a firm's hierarchy has been established as an important determinant of life-cycle wage growth, adding interest to the empirical investigation of promotion. Using personnel records from individual firms, both Baker, Gibbs, and Holmstrom (hereafter BGH, 1994a, 1994b) and Lazear (1992) demonstrated the importance of promotion on within-firm wage growth. Highlighting the importance of rising in level, BGH (1994b) found that levels alone explained about 70% of the variance in pay across employees in a given year. Lazear found that real wage declines were experienced by workers remaining more than seven years in the same job. Using a panel survey of households, McCue (1996) found that promotion accounted for nearly one-sixth of within-firm wage growth. The findings

relating to promotion in these papers and in a few others are described in more detail below.

BGH (1994a) examined 20 years of personnel data for all management employees of a single, medium-sized US firm in a service industry. Due to their interest in the effects of level on pay, it was necessary to identify the firm's hierarchical structure. They state that "Hierarchies are usually said to consist of job titles aggregated into levels related to the job's authority and place in the path of decision-making (hence the term level)." Because data on reporting relationships was unavailable, they relied exclusively on information about moves between job titles to define levels within the firm. Since there were large numbers of lengthy careers, characterized by movement through numerous job titles, they were able to clearly identify the hierarchical levels in the firm. Eight hierarchical levels and 17 major job titles encompassed over 99% of management level/salaried employment. In defining promotions, job title changes to higher levels were used.

Their findings on promotion included: a strong association between promotion and wage growth; the existence of promotion fast tracks in that employees promoted quickly at the low levels were promoted more often and more quickly later; the promotion probability was highest at low levels within the hierarchy (as in our data) which they attributed to a narrowing of the hierarchy; the promotion probability decreased with tenure in the firm; those promoted the fastest exited the firm more often; and new hires were initially promoted more quickly than incumbents but did not experience greater advancement over the course of their careers with the firm.[2]

BGH (1994b) also found substantial serial correlation in real wage growth for individuals, serial correlation that persisted even after observable differences between individuals were filtered out. This implied that heterogeneity across employees was explained only in part by differences in observable characteristics. Since strong wage growth was associated with more rapid promotion, BGH suggest that the presence of an unobserved variable, such as ability, drives both promotions and wage growth. This corresponds to our finding that unobserved heterogeneity is a central determinant of promotion.

Lazear (1992) examined 13 years of personal records to study the influence of job assignment on wages and turnover for full-time workers at a large durable-goods manufacturer. Higher level managerial employees were not included. Promotion was defined as a move to a job with a higher mean wage. Relating to promotion, individuals who change jobs tend to start with higher wages than those who do not and some jobs in the firm are more likely to lead to promotion than others. Job characteristics that are

associated with higher promotion rates include low average salaries and high average tenure and education. Production-related jobs also had higher promotion rates.

McCue (1996) studied 13 years of the Michigan Panel Study of Income Dynamics (PSID) to study the link between promotion and wage growth. Promotion in the PSID is one of the possible reasons for a job change. It is self-reported by survey respondents. McCue's findings include that better-paid workers were more likely to be promoted, that most promotions occurred early in a worker's career, and that promotions declined with time in position, experience, and age.[3]

An early investigation of promotion was conducted by Wise (1975). He examined a sample of 1,300 college graduates hired by a large US manufacturing firm, with a specified hierarchical structure, over the period from 1946 through 1964 to estimate the impact of individual characteristics on the promotion probability in the company. Promotion was defined as a move to a job with a higher grade level in the firm's hierarchical structure. The rate of promotion in grade level was found to be positively related to a set of personal characteristics including: college selectivity, college GPA, rank in graduate school, leadership ability, and initiative. The rate of promotion was negatively influenced by employee risk aversion (i.e., desire for job security). Unobserved individual personal characteristics were shown to have little effect on the promotion probabilities.

As well as summarizing the basic findings concerning internal labor markets in the literature, Gibbs and Hendricks (2004) analyzed five years of the personnel records of a large, US corporation. Each job change was classified by the firm as a promotion, demotion, lateral move, transfer, or exit. The firm adhered to fairly strict salary rules, with constraints enforced on workers near the top of salary ranges, but with promotions partially based on subjective performance ratings. Findings in regard to promotion included: support for "fast tracks," those promoted quickly the first time were more likely to be promoted again within three years; promotion rates that decrease with job and firm tenure; and promotions that occurred relatively more frequently at the bottom of salary ranges.

In addition to the papers discussed, there are other studies based upon samples of individual firms, including: Ariga, Ohkusa, and Brunello (1999), Chiappori, Salanic, and Valentin (1999), Seltzer and Merrett (2000), and Dohmen, Kriechel, and Pfann (2004). Evidence of promotion fast tracks was found also in Ariga, Ohkusa, and Brunello's single firm Japanese study, and in the Seltzer and Merrett's study of the Union Bank of Australia. Finally, promotions have been analyzed in the sociology and management

literatures.[4] Some findings in this literature include: (1) the positive influence of early promotions on later promotion;[5] (2) the positive influence of degree attainment on promotion;[6] (3) the importance of functional area and age in the attainment of a top executive position;[7] (4) lengthy firm tenure for top executives.[8] Promotion has also been examined in the context of gender discrimination.[9]

The remaining sections of the chapter are structured according to the following format. Section 2 describes the data. The static econometric model is introduced in Section 3 and its results are laid out in Section 4. Section 5 is devoted to examining how the effects of human capital, firm scale variables, and promotion opportunities on promotion probabilities vary across levels. A dynamic specification of the model that considers fast tracks is introduced in Section 6. Concluding remarks are found in Section 7.

2. DATA

The proprietary panel data set used in this study provides information on over 30,000 executives working at over 300 of the largest firms in the United States during the period from 1981 to 1988. Seventy percent of the firms in the sample are in manufacturing which includes, for instance, food, beverages, textiles, paper, chemicals, pharmaceuticals, glass, metal, machinery and electronic and transportation equipment. It was assembled by a major compensation consulting firm based on annual surveys completed by a human resource professional at the respondent company on both the company and individual executives. Respondent companies paid roughly $1,000 to participate in the survey, for which they received a report on the competitiveness of their pay levels relative to the pay levels of executives at a group of comparable firms. Firms were asked to complete the survey by the end of April and survey reports were to be distributed to respondents beginning in August of the survey year.[10]

The respondent company decided the number of executives to include each year and whether to participate annually or on a less frequent basis. The guidelines provided to firms suggested that they provide data on a representative sample of at least 75 executives in a variety of job families, managerial levels, and organizational units. When a job title was shared by many executives and firms did not wish to report on each, they were asked to report on several representative cases. Respondent companies submitting data on more than 120 executives in a given year were subject to an

additional fee. The mean number of executives reported on annually per firm was roughly 80 (Table A4).

The database reveals information on individual, job, and firm characteristics, including: age, years of education, functional area, job title, firm tenure, base pay, bonus pay, reporting level, industry, firm profits, sales, and employment. Gender is not available in these data. The consulting firm took measures to ensure that the information for each individual and company was valid and complete. All survey data were run through a series of error checking programs and subsequently staff reviewed for follow up with the respondent company when inconsistencies were noted. The information submitted on firm characteristics was accompanied by the respondent company's most recent annual report and proxy statement to ensure consistency of financial data.

A unique identifier assigned to each individual allows them to be tracked over time in their given firm. However, the movement of an individual between firms cannot be tracked as they would be assigned a new identifier in the subsequent company. An individual's disappearance from these data does not necessarily indicate an exit from the firm or a transition within the firm, as the respondent company elects which jobs to include each year.

One feature of these data is the classification of each executive into a reporting level in their company's hierarchy. The reporting level is the number of levels the position is located from the board of directors. The CEO is reporting level 1. All positions reporting directly to the CEO are assigned to reporting level 2; job titles at level 2 include the top legal officer, chief operating officer, and the top financial executive. All positions reporting directly to those at level 2 are assigned to reporting level 3; job titles at this level include profit center head, controller, top personnel executive, etc. Subsequent levels are defined in an identical fashion.

For use in our models of promotion probability, we define promotion as advancement in reporting level toward the CEO position. Promotions are not based on changes in job title. Hence, reporting levels can change without changes in job title, job titles can change without changes in reporting level, and they can change together. While Pergamit and Veum (1999) argue for a definition of promotion not limited to cases in which there is a change in job title, it is common in the promotion literature to do so. The benefits of a study across hundreds of firms in this case come at the cost of the ability to define hierarchical levels based on the observed job transitions possible in firm case studies, such as BGH, and may reduce the comparability of promotion rates across studies that take different approaches to defining

levels and promotions.[11] We intend to examine promotions based on job title changes in a subsequent chapter.

There are three advantages of the method used to define promotion in this chapter. First, the definition of promotion is consistent across firms. This does not mean that advancement in reporting level carries exactly the same implications for workers across firms. For instance, some firms may have flatter hierarchies in which promotions are more difficult to achieve and may carry greater rewards. Indeed, advancing a level within the same firm from different starting points in the hierarchy will not in general carry the same implications for wages gains and so forth. Even workers promoted one level within the same firm and from the same starting point will differ in their benefits from promotion depending on differences in age, expected tenure with the firm, and so forth. However, these considerations would be relevant to any method for defining promotion.

Second, our definition of promotion alleviates the need to construct firm hierarchies on some other basis. Reporting levels are self-reported by firms with the well-defined organizational structures common to Fortune 500 companies.

Third, to consider the existence of fast tracks, it is necessary to infer the executive's past speed of promotion on the basis of information available on the executive at the start of the sample. The eight-year sample period is too short to analyze fast tracks purely on the basis of promotions occurring within the sample. Making use of the executive's reporting level and age at the start of the sample is convenient for this purpose. The subsequent use of reporting level to define promotions that occur during the sample period provide a measure that is consistent with our measure of past speed of promotion.

In order to belong to our sample, executives had to meet the following two conditions: (1) only executives appearing in at least two consecutive years are kept for analysis because the construction of the promotion variable requires that we observe the executive's reporting level in two consecutive years to determine if a change took place (and only the consecutive years are used for executives with breaks in their reporting history); and (2) only executives who are first observed no more than five levels beneath the CEO are kept for analysis. There are 20,251 executives in the data for at least three years, 14,040 for at least four years, 8,766 for at least five years, 4,852 for at least six years, 2,900 for at least seven years, and 1,589 for eight years. Executives six or more levels from the CEO position are relatively few in number (see Table A3), constituting less than 6% of the sample of executives with at least two consecutive years of data. The average firm reports data on about six reporting levels (see Table A4). At levels far removed from the CEO position, the individuals come from a smaller set

of firms and may be less representative. For this reason, the analysis was cut off at six levels. The most common job titles among level 6 executives are plant manager, regional sales executive, and district sales executive.

Table A1 presents summary statistics for executives that have an initial observation at level 6 or at more senior levels in the firm. The table pertains to the executive's first year in the data. Table A2 presents summary statistics according to the observation number for executives of level 6 or higher. Executives are slightly more senior in level in later observations and have a lower incidence of promotion. Table A3 provides the fraction of executives promoted between their first and second years in the data by level. It also shows how tenure, age, and annual cash compensation (base pay and bonus) vary with level. These are recorded at the executive's first observation. It is clear from the table that the rate of promotion diminishes at higher hierarchical levels. For instance, 15% of level 5 executives were promoted between their first two observations. For level 3 executives, the promotion rate was only 4.3%. Age and annual cash compensation rise at more senior levels. Table A4 provides firm level statistics by year. On an annual average basis, firms include information on about 80 individuals and 6 reporting levels. The scale of the corporations in the sample is indicated by the profits, sales, and employment figures.

Table A5 provides a simple way of examining fast tracks through summary statistics. It examines whether younger executives have a current promotion advantage over older executives of the same level and firm. This is related to the notion of fast tracks because younger executives must have had more rapid advancement in their careers in order to have achieved the same level as older executives. In each executive's first two years in the data, there is slight promotion advantage for younger executives in levels 4, 5, and 6. There is no difference at levels 2 and 3.

3. THE STATIC ECONOMETRIC MODEL

In order to implement our empirical model of promotions, three issues should be addressed. These arise in both the static and the dynamic specification. The first issue relates to the identification of individual- and firm-specific unobserved characteristics. The second issue is the potential endogeneity of the initial rank observed for each individual in the panel. The third issue involves a standard choice that must be made when estimating dynamic discrete choice models, namely whether to use conditional maximum likelihood techniques (sometimes referred to as fixed effects

estimation), or unconditional inference techniques (random effects estimation). All of these issues deserve some discussion.

First, the distinction between individual- and firm-specific attributes is problematic, given the structure of the sample data. While it is possible to observe a few firm-specific variables (to be discussed below), the movement of executives between firms cannot be observed in the data set that we use. Therefore, the data do not allow us to separately identify the firm-specific unobserved term from the individual-specific term, unlike what is done in Abowd, Kramarz, and Margolies (1999). Without loss of generality, therefore, we refer to the unobserved factors as individual specific.

Second, it is conceivable that the initial level at which the individual is observed is affected by unobserved heterogeneity. In the panel data literature, this problem is referred to as the "initial condition problem." To address this issue, we need to either model the initial condition, or to model the distribution of unobserved heterogeneity conditional on the initial condition. In this chapter, we favor the latter option. As in Wooldridge (2005), we define the distribution of the unobserved heterogeneity term(s) conditional on the initial level. The heterogeneity term is decomposed into the sum of a regression component and an orthogonal unobserved component, both estimated flexibly (we use a finite mixture model).[12] This approach allows us to minimize the impact of distributional assumptions needed in order to implement such a model.

With respect to the third issue, our choice of an econometric estimation technique is largely dictated by the need to recover the marginal effects associated with the key variable, and allow for multi-dimensional population heterogeneity in promotion. It is also important to evaluate the relative importance of the determinants of promotion including human capital, unobserved heterogeneity, firm variables, and promotion opportunities. For these reasons, we focus on random effect estimation techniques, and treat unobserved heterogeneity as a random term, potentially correlated with strongly exogenous time-varying regressors.[13] While random effects techniques are often formulated in a context where the initial conditions of the stochastic process are modeled in a fully parametric framework, we propose a random effect estimation strategy based on flexible methods.

3.1. The Coding of Promotion

The aim of the model is to make inference about individual promotion histories from a sequence of rank levels (within a firm) occupied by

individuals. The sequence contains up to eight years of data. We define a promotion as a negative change in level (an accession to a higher rank in the hierarchy), that is

$$Y_{ijt} = 1(L_{ijt} - L_{ijt-1} < 0) \tag{1}$$

where L_{ijt} is the rank of individual i, in firm j, at time t and $1(\cdot)$ the indicator function.[14] In total, this results in seven potential promotion outcomes per individual. Promotion is coded as a binary variable. We do not distinguish between demotions and absence of promotions. A demotion, or an unchanged level, is recorded as a zero. Similarly, we do not make distinguish between one level promotions and promotions of more than one level – which are rare. A promotion of one, or more, level is coded as one. In order to minimize the impact of measurement error, if the level in the year subsequent to a promotion reflects a demotion, the original improvement in level is regarded as a coding error, and no promotion is recorded. This means that we do not code as promotions improvements in level that last just one year. Similarly, if a worker is a demoted in one year, but subsequently promoted in the next year to the original level, this return to the original level is not recorded as a promotion. Lastly, promotions are only registered when they constitute an improvement over the worker's initial level in the data. This ensures that promotions register only when they constitute a net upward movement over the span of observations, and not just upward movement over the previous year. Combined, these conditions on promotions reduce the original number of promotions from 11,620 to 8,489.

3.2. The Promotion Probability

The basic element of our econometric strategy is the following promotion probability

$$\Pr(Y_{ijt} = 1) = \Lambda(\beta_X \mathbf{X}_{it-1} + \beta_W \mathbf{W}_{jt-1} + \beta_{PO} PO_{ijt-1} + \beta_q L_{qit-1} + \alpha_i) \tag{2}$$

where q is the indicator of level (2, 3, 4, 5, 6).

- $\mathbf{X}_{it}$ represents a vector of individual-specific attributes, including years of education and age (both measured at the initial sample period), and tenure in the firm. Of these variables, only tenure is time varying. In job assignments models, the speed of promotion is positively related to the worker's innate ability and labor market experience.

- $\mathbf{W}_{jt-1}$ represents a vector of firm-specific time-varying variables, such as: firm size (employment), sales, and profits – all measured at $t-1$. Profits and sales are measured in millions of 1980 US dollars. Firm size is measured in thousands of employees. These variables are assumed to be exogenous in a strong sense.
- PO_{ijt-1} is an index between 0 and 10 indicating the fraction of employees in the level above the incumbent promoted in the current year. This is meant to measure the firm/level-specific density of promotion opportunities. It accounts for the aspect of promotion outcomes that are not driven by individual skills (observed or unobserved). If promotions depend on firm constraints, as well as individual capabilities, the impact of this variable should be high. However, to the extent that firm structure/hierarchy may be endogenous, the impact may be low. This variable is also assumed to be strongly exogenous.
- L_{qit-1} are endogenous time-varying binary indictors equal to 1 if the individual is at the rank level indicated by the subscript, and equal to 0 if not. Level 6 is the reference group. Level 1 (CEO) is not included because CEOs cannot be promoted internally.
- α_i is an individual-specific term which represents unobserved individual and firm heterogeneity. In order to resolve the initial condition problem, we specify the distribution of the unobserved ability term conditional on the initial level. However, to take into account that difference in initial levels are particularly meaningful after conditioning on age, we define a variable (agegap) measuring the mean age at the executive's level in his company minus the age of the individual. This variable captures the extent to which an individual has already achieved a rate of promotion higher than average. In order to take into account that the unobserved individual factor affecting promotion may not be orthogonal to time-varying variables (firm variables and promotion opportunities), we also allow for a relationship between heterogeneity and these same variables. More precisely, we define α_i as

$$\alpha_i = h_W(\bar{W}_j; \alpha_W) + h_{PO}(P\bar{O}; \alpha_{PO}) + h_{ag}(\text{agegap}_0; \alpha_{ag}) + \tilde{\alpha}_i \qquad (3)$$

  where $\bar{W}_{ij}$, $P\bar{O}_{ij}$, refer to the sample average (over the entire panel duration) of the time-varying regressors, where ag_0 is the initial agegap measured at date zero and where $h_W(\,)$, $h_{PO}(\,)$, and $h_{ag}(\,)$ are parametric functions specified as polynomials or order two.
- Our estimation method is based on the premise that $\tilde{\alpha}_i$ is characterized by an unknown cumulative distribution function, $H(\cdot)$, which is

approximated using a discrete distribution (Heckman & Singer, 1984). The type probability is

$$\Pr(\tilde{\alpha}_i = \alpha_k) = p_k$$

where $k = 1, \ldots, K$. The number of types, K, is assumed to be known, although it is the outcome of various experimentations. The type probability, p_k, is estimated using a logistic transform. In this chapter, we experimented with K ranging from two to four.[15]

- The promotion probabilities are assumed to be logistic, that is

$$\Lambda(.) = \frac{\exp(.)}{1 + \exp(.)} \tag{5}$$

Conditional on unobserved individual-specific heterogeneity, the promotion outcomes are assumed to be independent. This means that, given individual endowments, the promotion outcomes are random. Our empirical model is therefore more general than most job assignment models where promotion incidence is nonstochastic (see Eqs. (3) and (4)).

- $(\beta_x, \beta_W, \beta_q, \beta_{PO}, \alpha_W, \alpha_{PO}, \alpha_{ag})$ are parameters to be estimated and $\tilde{\alpha}_i$ is the residual individual-specific unobserved term with a distribution function that has to be estimated (approximated).

4. THE DETERMINANTS OF PROMOTIONS

In this section, we first present the parameter estimates obtained from the static model specification with unobserved heterogeneity in the intercept term. This specification ignores issues relating to fast tracks and promotion dynamics. In the most general case, we model the distribution of unobserved heterogeneity conditional on the "agegap" variable and allow the heterogeneity term to be correlated with all firm time-varying variables (profit, sales, and size) and the variable measuring promotion opportunities. We also report average promotion probabilities by level and types. Finally, using a variance decomposition of the index function, we illustrate the relative importance of the factors grouped into individual attributes, firm attributes, reporting level, and unobserved heterogeneity.

4.1. Parameter Estimates

As a first step, we experimented with the number of support points (K) for $\tilde{\alpha}_i$, examining estimations with up to four support points.[16] It turns out that the optimal specification requires two points of support. The results presented below are therefore for the case where $K = 2$. The parameter estimates are found in Table 1A. The first column is devoted to the model in which unobserved heterogeneity is allowed to depend on all firm variables as well as the agegap (model 1).

4.1.1. Education, Age, and Tenure

When education has a multiplicative effect on the growth in effective ability, the model's implication for the effect of education on promotion is clear. However, in the real world, we cannot rule out the possibility that more educated workers enter at more senior levels and that the effect of schooling is only located at the initial level. In the econometric model the effect of schooling is not restricted. Promotion probabilities are increasing in education (0.0421). The positive impact of education is consistent with the multiplicative assumption of schooling in effective ability and is robust to considerations of a variable level of entry.

Of the remaining human capital variables, only age is significant (-0.0244), tenure is not. Promotion probabilities are decreasing in age. The small magnitude of the tenure parameter is most likely explained by the relatively high dispersion in the intercept terms and suggests that, given unobservable factors, how long one has served in the firm is virtually irrelevant for the purpose of predicting promotion outcomes.

In order to illustrate the results, we present the marginal effects for the human capital variables and the firm variables (Table 1B). The marginal effects are computed for each individual and averaged over the entire sample. We report a standard deviation of the marginal effects which illustrates the cross-sectional differences in the marginal effects (for given parameter values). As logically expected from the small parameter values, the respective marginal effects are quantitatively unimportant (0.0024 for education, -0.00001 for tenure, and -0.0014 for age). As an example, four additional years of education has marginal effect on promotion probabilities in a given year of 0.0096.

In order to assess the relative importance of each variable (or group of variables), we decompose the single index function explaining the promotion probability. Our objective is to evaluate the explanatory power of each variable (or group of variables). Our measure of explanatory power

Table 1A. The Determinants of Promotions.

	Model 1	Model 2	Model 3
	Parameter (as t-ratios)	Parameter (as t-ratios)	Parameter (as t-ratios)
Individual human capital			
Education	0.0421	0.0413	0.0223
	(7.98)	(8.90)	(1.69)
Tenure	0.0001	− 0.0003	− 0.0016
	(0.55)	(0.81)	(0.47)
Age	− 0.0244	− 0.0267	− 0.0206
	(9.41)	(10.3)	(1.94)
Firm variables			
Promotion opportunities	0.7609	0.7441	0.7530
	(65.4)	(69.1)	(53.4)
Firm profits	0.0148	0.0115	− 0.0970
	(0.66)	(0.45)	(3.36)
Firm sales	0.1603	0.1505	0.1143
	(3.28)	(7.66)	(4.77)
Firm size	0.0239	0.0271	− 0.0262
	(2.83)	(2.12)	(1.94)
Level in the firm			
Level 6	–	–	–
Level 5	− 0.2581	− 0.2454	− 0.3454
	(7.56)	(6.91)	(14.9)
Level 4	− 0.6523	− 0.6216	− 0.5209
	(18.8)	(17.8)	(12.9)
Level 3	− 1.2186	− 1.1717	− 1.2167
	(34.1)	(38.9)	(17.2)
Level 2	− 7.2911	− 7.5005	− 7.3034
	(32.2)	(32.0)	(86.9)
Unobserved heterogeneity			
$\tilde{\alpha}_{type1}$	− 4.9701	− 5.1097	− 5.0119
	(44.57)	(47.63)	(50.76)
$\tilde{\alpha}_{type2}$	− 1.5325	− 1.4254	− 1.5560
	(71.93)	(54.44)	(56.94)
Initial agegap	0.0086	0.0084	0.0085
	(3.04)	(6.46)	(0.10)
Initial agegap^2	− 0.0005	− 0.0006	− 0.0004
	(9.57)	(10.71)	(4.95)
Mean opportunities	− 0.0727	–	–
	(2.62)		
Mean opportunities^2	− 0.0343	–	–
	(2.91)		

Table 1A. (*Continued*)

	Model 1	Model 2	Model 3
	Parameter (as *t*-ratios)	Parameter (as *t*-ratios)	Parameter (as *t*-ratios)
Mean sales	−0.1205	−0.1171	–
	(6.16)	(5.04)	
Mean sales^2	−0.0018	−0.0016	–
	(3.95)	(3.40)	
Mean profits	−0.1001	−0.0861	–
	(2.04)	(2.45)	
Mean profits^2	0.0036	0.0031	–
	(2.70)	(2.19)	
Mean size	−0.0358	−0.0353	–
	(2.83)	(2.90)	
Mean size^2	0.0005	0.0004	–
	(1.03)	(0.85)	
Pr(type 1)	0.3543	0.3728	0.4841
	(15.25)	(15.38)	(20.03)
Mean log likelihood	−0.627558	−0.628041	−0.628472

is the percentage loss in the explanatory power of the index function regression when a variable or group of variables is omitted. The results in Table 1C illustrate the relative unimportance of the human capital variables. Omitting human capital variables (age, tenure, and education) reduces explanatory power by only 1%.

4.1.2. Firm Attributes: Promotion Opportunities, Profits, Sales, and Size

In our analysis, the promotion opportunity index is meant to capture the sources of promotion outcomes that are not driven by individual skills (observed or unobserved). With regards to promotion opportunities, to the extent that firm hierarchies are rigid and individual promotion outcomes are not independent of vacancies in the hierarchy, we should expect promotion opportunities to play a critical role in determining promotion probabilities. The estimate for the effect of the promotion index is estimated to be 0.7609 (Table 1A). Given that the index is a number between 0 and 10, this implies that opportunities are relatively important. This is verified by examining the related marginal effect which is equal to 0.0438 (Table 1B). The estimate indicates that if everyone observed the year before and one level above is

Table 1B. Some Marginal Effects.

	Model 1	Model 2	Model 3
	Estimate (SD)	Estimate (SD)	Estimate (SD)
Human capital variables			
Education	0.0024	0.0024	0.0011
	(0.0025)	(0.0026)	(0.0113)
Tenure	− 0.00001	− 0.00001	− 0.0001
	(0.00001)	(0.00001)	(0.0001)
Age	− 0.0014	− 0.0014	− 0.0010
	(0.0015)	(0.0015)	(0.0012)
Firm variables			
Promotion opportunities	0.0438	0.0439	0.0369
	(0.0463)	(0.0470)	(0.0456)
Profits	0.0008	0.0007	− 0.0047
	(0.0009)	(0.0006)	(0.0059)
Sales	0.0092	0.0088	0.0056
	(0.0098)	(0.0095)	(0.0008)
Size	0.0014	0.0016	− 0.0013
	(0.0014)	(0.0017)	(0.0016)

Note: The marginal effects are averaged over all individuals. The reported standard deviations is a measure of cross-sectional dispersion in the marginal effects, given parameter estimates.

Table 1C. Variance Decomposition of the Index Function: The Loss in Explanatory Power for Each Group of Variables.

Variables	1	2	3
Human capital (age, tenure, education)	1%	1%	0.5%
Firm variables (profit, sales, size)	10%	8%	2%
Promotion opportunities	18%	15%	16%
Level in the firm	34%	31%	30%
Unobserved heterogeneity	42%	41%	40%

Note: The percentages denote the loss in explanatory power of the explained part of the index function regression for each group of variables. They are computed from the difference in the coefficient of correlation from the regression that includes all factor and a regression that excludes only each particular group.

promoted in the current year, promotion probabilities for executives in the level beneath would increase about 0.44. The variance decomposition exercise in Table 1C indicates that excluding promotion opportunities reduces explanatory power about 18% in model 1.

The profits, sales, and size (employment) are included in order to have additional control variables for unmeasured firm factors as firm unobserved heterogeneity cannot be distinguished from individual unobserved heterogeneity. These controls may be useful because Eriksson and Werwatz (2005) found a slight positive association between the rate of promotion and firm size. We find that the effect of firm profit (0.0148), sales (0.1603), and size (0.0239) on promotion are positive (Table 1A). The related marginal effects are however quite small (Table 1B). They may indicate that the promotion process of American executives, after accounting for promotion opportunities, is not sensitive to the business cycle. As a group, excluding firm profit, sales, and size reduces explanatory power by 10% (Table 1C).

4.1.3. Differences in Level

There is no clear theoretical basis on which to attach a sign on the effect of level on promotion after conditioning on all individual and firm attributes. In job assignment models, though both the threshold level of effective ability that must be achieved for promotion and the average level of effective ability among those at a given level increase with level, the relative increase is not specified. In tournament models, the promotion probabilities depend on exogenous parameters such as the number of competitors in relation to the number of slots available. These are typically exogenously determined.

We find that promotion outcomes are substantially dependent on the current level of the manager. The level-specific dummies (ranging from -0.2581 at level 5 to -7.2911 at level 2) indicate that given all individual- and firm-specific endowments, promotion probabilities become smaller as one reaches higher ranks (Table 1A). In particular, the promotion probability approaches zero when individuals reach level 2.

As the level variable is discrete, the marginal effects are illustrated by the differences in promotion probabilities across levels found in column 1 of Table 1D. They indicate that, although the average annual promotion probability is 0.0861, the level-specific average probabilities range from 0.1213 (level 6) to 0.0004 (level 2). The average promotion probability is reduced by 0.0302 as one moves one level up in the hierarchy. The variance decomposition in Table 1C indicates that excluding reporting level reduces explanatory power by about 34%.[17]

4.1.4. Unobserved Heterogeneity

In our econometric model, unobserved heterogeneity refers to individual-specific attributes that are persistent over the life cycle. By definition, we do not know the extent to which these factors represent factors unknown to the

Table 1D. Promotion Probabilities by Level and Types in Model 1.

Rank	Probabilities		
	Population average	Type 1	Type 2
Level 6	0.1213	0.0151	0.1787
Level 5	0.1014	0.0125	0.1496
Level 4	0.0771	0.0093	0.1139
Level 3	0.0513	0.0060	0.0758
Level 2	0.0004	0.0000	0.0006
All levels	0.0861	0.0116	0.1261

Note: The promotion probabilities are averaged over all individuals at a particular rank.

current firm or the outside labor market. Therefore, there is no one-to-one correspondence between unobserved heterogeneity (from our perspective) and the complement of the information set of the firm, or the outside labor market.

The estimates for the type-specific intercept terms (the $\tilde{\alpha}_k$) in Table 1A indicate an important dispersion across types (-4.9701 vs. -1.5325). It follows that type 1 individuals, who correspond to 35% of the population, are rarely promoted, while type 2 individuals are much more likely to experience promotions. The type-specific promotion probabilities, found in the last two columns of Table 1D, indicate the importance of the dispersion. Obviously, this dispersion implies that, over a sustained period, the stochastic process generating promotion will depict serial correlation in promotions. In the literature, this is referred to as the fast track hypothesis. As such, this high degree of heterogeneity indicates the presence of serial correlation, but is not sufficient to establish whether or not there is a causal fast track.

Except for the initial agegap variable, which measures the mean age at the executive's level less the executive's age and represents a relative rate of promotion speed achieved by the start of the panel, the regression components of the distribution unobserved heterogeneity are difficult to interpret – as theory does not offer guidance. High promotability appears to be positively, but weakly, correlated with the initial speed of promotion – as measured by the agegap variable. The effect is concave.

Finally, the importance of unobserved heterogeneity is well illustrated in the variance decomposition. It accounts for 42% of the total variation, and it is the most important component of the stochastic process determining promotion (Table 1C). At this point we can address an important general

question in regards to careers, the extent to which it is the individual human capital (education, tenure, age, and some portion of unobserved heterogeneity) versus the environment (promotion opportunities that vary by firm and level, profit, sales and size that vary by firm, reporting level that is modeled to have a common effect across firms, and some portion of unobserved heterogeneity) that determines promotion. The effect of level is independent of the unobserved heterogeneity term and provides an estimate of the exogenous effect of level change. It must therefore be grouped with the firm variables. Unobserved heterogeneity is not exclusively due to individual heterogeneity. As explained previously, we cannot separately estimate firm and individual effects. Therefore, depending on unobserved heterogeneity, we can say that the removal of factors related to the individual reduces the explanatory power of the index function by at most 43% (in the case where all omitted heterogeneity would be individual specific) and at least 1% (if all heterogeneity would be firm specific (Table 1C)). While this is a wide range, it can be said that less than half of the promotion process is driven by individual characteristics. This points to the importance of considering the influence of the environment in theoretical models of promotion.

4.1.5. Additional Specifications

The second column of Table 1A (model 2) contains estimates for the case where the correlation between heterogeneity and promotion opportunities is forced to be zero. In the third column (model 3), the unobserved heterogeneity term only depends on the agegap variable. Because conditioning on the entire path of the time-varying variables requires strong exogeneity, these alternative approaches are important in order to establish the robustness of the results. The results appear relatively robust. This may be seen upon looking at the corresponding marginal effects and at the variance decomposition. While the third specification implies a much reduced role for firm variables, the rankings of the different groups of variable does not change (i.e., unobserved heterogeneity, level, and promotion opportunities are the three most significant determinants of promotion).

5. THE EFFECT OF HUMAN CAPITAL AND PROMOTION OPPORTUNITIES ACROSS LEVELS

Until now, we have focused on a model specification that forces separability between promotion opportunities, unobserved ability, and the individual

level occupied in the firm. This assumption may be quite restrictive and it may be relaxed by allowing the effect of all variables (as well as unobserved heterogeneity) to depend on level. With these changes, the model becomes

$$\Pr(Y_{ijqt} = 1) = \Lambda(\beta_{qW} W_{jt-1} + \beta_{qX} X_{it-1} + \beta_{qPO} PO_{ijt-1} + \alpha_i \beta_{q\alpha}) \qquad (6)$$

for $q = h$ and l. The distinction between h and l is meant to capture differences in promotion outcomes at higher levels (h refers to levels 2 and 3) and lower levels (l refers to levels 4, 5, and 6) and is dictated by the desire to obtain a tractable number of parameters and to facilitate comparisons between high and low levels in the hierarchy.[18] For clarity, we use the following parametrization:

$$\beta_{h,s} = \tilde{\beta}_{hs} \beta_{l,s} \qquad (7)$$

for $s = W, X, PO$, and α and we normalize $\beta_{l\alpha}$ to 1.

These estimates are found in Table 2A. For each variable, we report both $\beta_{l,s}$ and $\tilde{\beta}_{hs}$. The estimate corresponding to $\tilde{\beta}_{hs}$ indicates if the parameter estimate for the effect of the specific variable increases at higher levels (when it exceeds 1) or decreases (when it is below 1). However, it should be noted that because we no longer have a level-specific intercept term, the decreasing effect of level on promotion incidence is captured in β_hs. For this reason, most of the β_hs are below one. The results indicate that the parameters have not changed sign (except for firm profits). However, a change in the relative importance of the variables has taken place and will be explained by the relative decrease in the respective β_hs.

To get a more clear picture, we have performed a variance decomposition of the index function at high and low levels. These are found in Table 2B. There are two major highlights. First, promotion opportunities have much less explanatory power in the promotion of high-level executives. Its explanatory power drops from 71% at low levels in the hierarchy to 12% at high levels. This suggests that individual promotion outcomes at the top levels in the firm are less determined by vacancies. The great majority of promotions for high-level executives, levels 2 and 3 as classified in this estimation, are from level 3 to 2, and not from level 2 to CEO. Second, unobserved heterogeneity has much more explanatory power for high-level executives. Its explanatory power goes from 23% at low levels to 84% at high levels.

Table 2A. Models with Interactions: The Effects of Human Capital and Promotion Opportunities by Level.

		Parameter	(SE)
Individual human capital			
Education	β_l (level 4, 5, 6)	0.0400	(0.0033)
	$\tilde{\beta}_h$ (level 2, 3)	0.6602	(0.0286)
Tenure	β_l (level 4, 5, 6)	0.0007	(0.0020)
	$\tilde{\beta}_h$ (level 2, 3)	0.9615	(0.0029)
Age	β_l (level 4, 5, 6)	−0.0272	(0.0012)
	$\tilde{\beta}_h$ (level 2, 3)	1.4157	(0.0074)
Firm variables			
Promotion opportunities	β_l (level 4, 5, 6)	0.7542	(0.0125)
	$\tilde{\beta}_h$ (level 2, 3)	0.4421	(0.0189)
Firm profits	β_l (level 4, 5, 6)	−0.0259	(0.0120)
	$\tilde{\beta}_h$ (level 2, 3)	0.9672	(0.0016)
Firm sales	β_l (level 4, 5, 6)	0.0111	(0.0040)
	$\tilde{\beta}_h$ (level 2, 3)	0.9852	(0.0007)
Firm size	β_l (level 4, 5, 6)	0.0755	(0.0180)
	$\tilde{\beta}_h$ (level 2, 3)	0.9367	(0.0034)

Unobserved Heterogeneity	Parameter	(SE)
α^S_{0type1}	−2.3451	(0.0476)
α^S_{0type2}	−2.4067	(0.0368)
Initial agegap	0.0382	(0.0031)
Initial agegap^2	−0.0011	(0.0000)
Mean promotion opportunities	−0.0613	(0.0049)
Mean promotion opportunities^2	−0.0587	(0.0080)
Mean sales	0.0524	(0.0135)
Mean sales^2	−0.0015	(0.0004)
Mean profits	−0.0799	(0.0068)
Mean profits^2	0.0047	(0.0012)
Mean size	−0.1067	(0.0217)
Mean size^2	0.0012	(0.0005)
$\tilde{\beta}_l$ (level 4, 5, 6)	1.0000	
$\tilde{\beta}_h$	0.8945	(0.0122)
Pr(type 1)	0.5653	(0.0381)
Mean log likelihood	−0.644236	

Table 2B. Variance Decomposition of the Index Function: The Loss in Explanatory Power for Variable Groups by Level.

Variables	Level 4, 5, 6	Level 2, 3
Human capital (age, tenure, schooling)	4%	5%
Promotion opportunities	71%	12%
Firm variables (profits, sales, size)	6%	4%
Unobserved heterogeneity	23%	84%

Note: The percentages denote the loss in explanatory power of the explained part of the index function regression for each group of variables. They are computed from the difference in the coefficient of correlation from the regression that includes all factor and a regression that excludes only each particular group.

6. PROMOTION DYNAMICS AND FAST TRACKS

In the dynamic specification, it is necessary to efficiently summarize past promotion histories. As a starting point, we consider the estimation of a general dynamic promotion probability model, which, ideally, would take the following form

$$\Pr(Y_{ijt} = 1) = F(\varpi_{it}, \varpi_{jt}, Y_{ijt-1}, Y_{ij,t-2}, \ldots, Y_{ijt-p}, L_{ijt_0})$$

In this expression, ϖ_{it} and ϖ_{jt} are individual- and firm-relevant attributes, $(Y_{ijt-1}, Y_{ij,t-2}, \ldots, Y_{ijt-p})$ is a p dimensional vector of relevant past promotion outcomes, and L_{ijt_0} is the starting level (at time t_0). At this stage, the relevant question is how to summarize the entire vector of past promotion histories in a reasonable way. In the econometric literature devoted to the estimation of dynamic logit models with fixed effects (Chamberlain, 1984; Magnac, 2000), it is pointed out that nonparametric identification of two lags requires at least seven periods. It is reasonable to expect that the effect of past promotion goes beyond lags of order two or three. For this reason, we disregard the short run dimension of promotion dynamics, and focus on a summary of all past promotion outcomes.[19]

Ideally, we would like a measure of past promotion history that embodies the signal provided to the labor market regarding the caliber of the executive. The theoretical literature considers the importance of the signal provided by initial promotion, assuming a common starting level. Were the initial placement levels considered, as well as promotions, the importance of this signal would be just as relevant. Our measure should therefore be to capture the effect of early promotion history as well as the level of the initial

placement in the firm's hierarchy. In order to capture both aspects, we define a speed of promotion variable (referred to as *Speed* below) which is measured as the ratio of the level an executive has risen to by the start of the sample to the executive's years of labor market experience. With level 1 representing CEOs, levels fall with promotions and higher initial assignments. Since it is intuitively easier to think of promotion speed as a positive number, we look at the level an executive has risen to at the start of the sample in reference to an arbitrarily selected reference level 12. The reference level chosen is irrelevant, as it changes the number of levels an executive has risen to equally across executives. If we were to measure the speed of promotion only by considering the number of promotions, those who entered the labor market at a higher level would have fewer promotions due to starting closer to the top of their hierarchies. As such, we would then be confounding these executives with those who started beneath, and have a lower promotion probability for other reasons.

6.1. Modeling the Fast Track Hypothesis

In the dynamic setting, the promotion probability is now

$$\Pr(Y_{ijt} = 1) = \Lambda(\beta_X X_{it-1} + \beta_W W_{jt-1} + \beta_{PO} PO_{ijt-1} + \beta_q L_{qit-1} + \beta_i Speed_{it-1}) \tag{8}$$

where the variable $Speed_{it}$ measures the speed of promotion achieved up to date t. It is calculated as the ratio of the number of levels (#*levels*) reached at any point in time (in reference to an assumed starting level of level 12) and the difference between age and years of education (minus five). It is meant to capture the causal fast track hypothesis. As individuals are observed over the sampling period, the speed of promotion is adjusted according to the following law of motion

$$Speed_{it} = \frac{\#Levels_{t-1} + Y_{it}}{(Age_{t-1} - Education - 5) + 1} \tag{9}$$

In the most general version, we allow for heterogeneity in the effect of fast track. To do this, we allow for interaction terms between individual attributes and the speed of promotion variable. We pay a particular attention to two sets of variables, observed human capital (education and tenure) and differences in levels. Given that the initial speed of promotion achieved at the start is also correlated with heterogeneity, the individual-specific slopes are

expressed as

$$\beta_i = h_W(\bar{W}_j) + h_{PO}(P\bar{O}) + h_{ag}(ag_0) + h_S(Speed_0) + \beta_{S1} Education_i + \beta_{S2} Tenure_{it} + \beta_{S3} Level_{it} + \tilde{\beta}_i \quad (10)$$

where the $h_s(\cdot)$ are parametric functions as defined in Eq. (7).

While there might exist more flexible methods to allow for interactions (such as spline functions allowing for the slope to differ at some of the possible values of education and tenure), we retain the standard interaction term in order to keep the number of parameters at a manageable level and because our objective is only to infer the sign of the derivative of the slope with respect to tenure and education.

In total, we report three sets of estimates. These estimates are found in Table 3A. In the first set, in the first column, the effect of past speed of promotion is summarized in two "type-specific" parameters. In the second column, we allow the effect of the speed of promotion to depend on observed human capital (education and tenure). Finally, the results reported in the third column allow us to investigate how the effect of past speed of promotion changes as individuals move up in the hierarchy. Given the focus of this section, the most interesting estimates pertain to the effects of the speed of promotion on promotion outcomes.

6.2. Is there a Causal Fast Track?

As was the case for the promotion probability intercept terms earlier, there is also substantial heterogeneity in the unobserved parts of the slope parameters $\tilde{\beta}_i$. The regression function that relates the slopes to endogenous initial conditions indicate that those who have a higher slope tend to have achieved a higher speed of promotion at the start of the sampling period. Although the effect of speed of promotion will be positive for some individuals, but negative for others (depending on the initial speed and the promotion opportunities), the negative effect will dominate for the average individual. Indeed, the averages and standard deviations for the slope (along with the minima and maxima) found in Table 3B, indicate that the effect is negative on average (-0.5523), although past promotion speed is positive for a subset of the population. This implies that the potential comparative advantage earned by those who have been promoted earlier, referred to as the causal fast track, is very weak, and quantitatively unimportant.[20] This finding suggests that job assignment models producing fast tracks

Table 3A. Investigating the Causal Fast Track Hypothesis.

	Model 1	Model 2	Model 3
	Parameter (as *t*-ratios)	Parameter (as *t*-ratios)	Parameter (as *t*-ratios)
Individual human capital			
Education	0.0866	0.1226	0.0891
	(7.98)	(13.3)	(10.8)
Tenure	− 0.0020	0.0079	− 0.0019
	(1.07)	(1.55)	(1.06)
Firm variables			
Firm profits	− 0.0450	− 0.0454	− 0.0473
	(3.37)	(3.49)	(3.86)
Firm sales	0.0175	0.0176	0.0188
	(2.47)	(2.53)	(2.80)
Firm size	− 0.0049	− 0.0045	− 0.0060
	(1.14)	(1.07)	(1.41)
Promotion opportunities	0.6713	0.6702	0.6609
	(65.1)	(65.8)	(74.1)
Level in the firm			
Level 6	–	–	–
Level 5	− 0.1487	0.3278	− 0.1503
	(3.88)	(3.79)	(5.15)
Level 4	− 0.5092	− 0.0499	− 0.3445
	(12.7)	(5.53)	(12.3)
Level 3	− 1.1263	− 0.1543	− 0.8830
	(22.2)	(0.04)	(26.4)
Level 2	− 6.8773	− 0.0414	− 5.7968
	(26.3)	(8.56)	(29.1)
Unobserved heterogeneity			
$speed_{type1}$	− 0.1820	0.0336	− 0.2791
	(24.3)	(3.84)	(33.2)
$speed_{type2}$	− 0.8473	0.0872	− 0.8352
	(22.5)	(3.59)	(33.6)
$speed \cdot educ$	–	− 0.0411	–
		(6.12)	
$speed \cdot tenure$	–	− 0.0371	–
		(2.04)	
$speed \cdot level5$	–	–	− 0.1188
			(8.90)
$speed \cdot level4$	–	–	− 0.5955
			(31.0)
$speed \cdot level3$	–	–	− 0.7989
			(35.5)

Table 3A. (*Continued*)

	Model 1	Model 2	Model 3
	Parameter (as t-ratios)	Parameter (as t-ratios)	Parameter (as t-ratios)
speed · *level*2	–	–	− 1.7893 (29.2)
Pr(*type* 1)	0.5523 (10.3)	0.5026 (9.36)	0.5002 (8.28)
Initial speed	0.0408 (10.2)	0.0653 (23.8)	0.0010 (0.82)
Mean promotion opportunities	− 0.1752 (2.85)	− 0.1907 (3.03)	− 0.1404 (2.54)
Mean log likelihood	− 0.628297	− 0.628243	− 0.628375

Table 3B. The Distribution of the Effect of Past Speed of Promotion.

	Model 1	Model 2	Model 3
Mean	− 0.5523	− 1.1874	− 0.8634
SD	0.3901	0.7428	0.6047
Minimum	− 2.5799	− 3.3941	− 3.0234
Maximum	0.3212	0.1524	0.2790

Note: The effect of past speed of promotion is computed from $\beta_i = h_W(\bar{W}_j) + h_{PO}(P\bar{O}) + h_{ag}(ag_0) + h_S(Speed_0) + \beta_{S1} Education_i + \beta_{S2} Tenure_{it} + \beta_{S3} Level_{it} + \bar{\beta}_i$. The average are taken over all individuals and types.

solely due to worker heterogeneity (models with full information or symmetric learning) are capturing the primary source of serial correlation in promotion outcomes.[21]

The negativity of the promotion dynamics parameter is interesting in its own right, and deserves some attention (Table 3B). Negative fast track effects are, as far as we know, never mentioned in the literature. Our reading of the fast track effect is that the speed of promotion raises current and subsequent promotion probabilities, those who have been promoted first will build a comparative advantage in promotions. In practice, negative fast track effects may take the following form. In a world in which individual abilities are eventually known (the assumption under symmetric learning in the job assignment literature), and where identical individuals achieve the

same final level, the realization of an abnormally high rate of early career promotions may simply be compensated by a lower promotion rate later.

6.3. The Effects of Education and Tenure on Promotion Dynamics

If more educated workers are advantaged in promotion, it is possible that their past promotion histories are less important and that differences in education might account for a portion of the cross-sectional differences in individual-specific slopes. In the job assignment model with asymmetric information, the promotion of less educated workers provides a signal to the outside labor market of exceptional ability. The promotion of more educated workers sends a weaker signal of ability as they are already viewed from the outside as more able. In the event that past promotion histories are used as a signal, the significance of the signal is therefore decreasing with schooling. The relevant estimates are found in the second column of Table 3A. As conjectured, the estimate for the interaction term between education and speed of promotion is negative (-0.0411). It indicates that, as an individual gets more schooling, the effect of past promotion goes toward zero (or negative numbers).[22]

If past promotion histories play a signaling role, it is natural to expect this role to diminish with accumulated labor market experience or tenure. Uncertainty in regards to worker ability should be at its greatest upon a worker's entry into the labor force and the signals provided by the early career history should have the most impact on the beliefs about workers with less seniority. This assertion is consistent with the results of model 2. The parameter estimate for the interaction term between tenure and speed of promotion is negative (-0.0371). The findings in regards to education and tenure seem consistent with job assignment models with asymmetric learning.

6.4. The Effect of Past Promotions by Hierarchical Level

Difference in levels may also account for a certain degree of heterogeneity in the effect of past promotion histories. At the theoretical level, it is also conceivable that the signaling aspect of past promotions may vanish as an executive reaches higher level. The estimates found in model 3 (Table 3A) indicate that the effect of fast track decreases uniformly from lower levels (level 6) to higher levels (level 2). The estimates range from -0.1188 (level 5) to -1.7893 (level 2). When averaged over all individuals and types,

the average parameter is negative (-0.8634) and, again, it is positive for a subset of the population (see Table 3B).

To summarize, on average, causal fast track effects are qualitatively small and are not a key determinant of observed promotion histories. Nevertheless, it is interesting that the evidence suggests that causal fast tracks are associated with lower level workers with less education and tenure which provides some support for the promotion signaling hypothesis.

6.5. The Relative Importance of the Causal Fast Track

Finally, to quantify the relative importance of the speed of promotion variable, we have performed a variance decomposition. Since the dynamic model is specified differently than the static version (for instance unobserved heterogeneity is multiplicative in the promotion speed variable), the explanatory power percentages reported in Table 3C may differ from those reported previously in Table 1C. The key finding in Table 3C is that the effect of past speed of promotion is quantitatively small. The past speed of promotion accounts for 1% or less of the index function. This is the case for all three model specifications and it means, for the sake of a comparison, that past speed of promotion is no more important than standard human capital variables. In practice, this also suggests that promotion process may be modeled as a static discrete choice model.

Table 3C. Variance Decomposition of the Index Function: The Loss in Explanatory Power for Each Group of Variables.

Variables	1	2	3
Human capital (age, tenure, education)	1%	1%	2%
Firm variables (profit, sales, size)	2%	3%	1%
Speed of promotion	1%	0.5%	1%
Promotion opportunities	26%	34%	41%
Level in the firm	52%	46%	50%
Unobserved heterogeneity	20%	17%	28%

Note: The percentages denote the loss in explanatory power of the explained part of the index function regression for each group of variables. They are computed from the difference in the coefficient of correlation from the regression that includes all factor and a regression that excludes only each particular group.

7. CONCLUSION

In this chapter, a model of the promotion process is fit on a panel of American executives. The model allows promotion probabilities to depend on endogenous past promotion histories, observable human capital endowments, time-varying firm-specific variables, and unobservable heterogeneity. The results shed light on the complex process that governs hierarchical transitions. The stochastic process that drives promotion might be thought of as a series of probabilities that are smaller for individuals higher in the hierarchy. These probabilities are largely dependent on unobservable individual heterogeneity, level, and promotion opportunities. The standard human capital variables (age, tenure, and schooling) are virtually unimportant in predicting current promotion probabilities for executives after controlling for endogenous initial conditions and unobserved individual heterogeneity. Even if all unobservable heterogeneity is attributed to the worker (a portion reflects firm heterogeneity), then less than half of the promotion process is determined by individual factors. Models neglecting to consider the role of promotion opportunities and promotions rates that vary with levels are missing important aspects of the promotion process.

Fast tracks in the job assignment literature have been motivated by differences in ability in both full information and asymmetric learning models, and by promotion signaling in models with asymmetric learning. The econometric model distinguishes between the extent to which fast tracks arise out of heterogeneity and out of the causal advantage gained by early promotion. Our results indicate that the fast tracks documented in the empirical literature result largely from unobserved individual heterogeneity and not from rapid early promotions having their own inherent effect on later promotions. Therefore, fast tracks are largely explained without an appeal to promotion signaling.

Going beyond this overall result in regard to fast tracks, there is evidence of high cross-sectional dispersion in the effect of past promotion histories on promotion probabilities. For most individuals we find evidence of a negative correlation between the effect of past promotion rates on current promotion probabilities. However, for a minority of the population the effect is positive. Moreover, the individual-specific effect of achieving a high rate of past promotion on promotion probabilities is negatively related to education and tenure. Consistent with the asymmetric information hypothesis that the signaling aspect of past promotion is stronger for those who are less educated or new to the firm, structural fast tracks are stronger

for individuals with less education and tenure. This is an important empirical finding in support of the notion of job assignment signaling.

In future work, we intend to examine the role of the executive's functional area on promotion probabilities and the robustness of these results to an alternative definition of promotion based on job title change. It would also be valuable to investigate the relative importance of human capital and endogenous promotions in explaining lifetime earnings and the nature of serial correlation in wage growth.

NOTES

1. See Heckman (1981) for a general discussion of true versus spurious state dependence.

2. Treble, van Gameren, Bridges, and Barny (2001) replicated the BGH (1994a) analysis using nine years of personnel data from a large British financial sector firm with a specified hierarchical structure containing seven management levels. Promotion was defined as a job transition to a higher grade level in the firm's hierarchical structure. For comparison with our chapter, for employees in management levels, Treble et al. found promotion rates varied negatively with level and that fast track promotion effects and fast track "exit" effects existed (those promoted more quickly tended to have a higher exit rate).

3. Others papers utilizing multi-firm data to study promotion include Eriksson and Werwatz (2005), DeVaro (2006), and DeVaro and Brookshire (2007). Eriksson and Werwatz examined a sample of 222 Danish firms and constructed job levels based on the broad job classifications that they specified (1, top managers; 2, high-level managers; 3, middle management and supervisors; 4, nonmanagerial white-collar workers and skilled blue-collar workers; 5, unskilled blue-collar workers; and 6, other workers). Promotion was defined as a movement into a higher job level. Interestingly, at their broad aggregation of jobs into job levels, promotions from within the firm were not a prominent feature in most of their firms, though higher rates of promotion and longer careers characterized the finance and utilities industries. Additionally, the incumbent status of a worker promoted to a given level from within the firm did not increase the probability of future promotions over that of newly hired employee.

DeVaro (2006) and DeVaro and Brookshire (2007) utilized the Multi-City Study of Urban Inequality (MCSUI), a cross-sectional employer telephone survey of over 3,000 establishments collected between 1992 and 1995. The survey questions pertaining to the establishment's most recently hired worker form the basis of the analysis. Promotion was defined based on the firm's response to the question as to whether this worker had received a promotion by the survey date or whether a promotion was expected in the next five years. Relating to promotion, DeVaro's results suggest that both actual promotions and expected promotions are associated with higher relative performance and for-profit status. DeVaro and Brookshire

(2007) found that workers are less likely to receive promotions in nonprofit organizations and nonprofits were less likely to base promotions on job performance.

4. For surveys see Forbes and Piercy (1991) and Rosenbaum (1984).

5. Rosenbaum (1979) finds that those promoted first were more likely to receive further promotions and reach higher levels in the firm. Howard and Bray (1988) find that Bell System managers with more significant job challenges in years 1 through 8 exhibited greater advancement at year 20.

6. Howard and Bray (1988) find a college degree to be the best predictor of promotion. Forbes and Piercy (1991, p. 165) find that the time to the CEO position is reduced through higher levels of education. Useem and Karabel (1986) show the importance of earning a degree from an elite institution when the executive is not from elite social origins.

7. Vroom and MacCrimmon (1968) find that promotion opportunities vary with functional area and are better in finance and marketing. Forbes and Piercy (1991, p. 4) find the functional area backgrounds of CEOs to vary by industry. They also find with regards to eventual CEOs that the time to reach various top positions in the organization varied by functional area (p. 145) and provided evidence of age varying systematically with career level (p. 144). For example, CEOs reach a top management position by age 47 on average and none reach this level after age 58. Out of 230 CEOs, none were promoted to the CEO position later than age 65, the mean age was 50.

8. Forbes and Piercy (1991, p. 5) note that successful top executives spend most of their careers within the same firm. Tuckel and Siegel (1983) find most CEOs to have spent their entire careers within one firm.

9. A few findings in this literature include: that women are crowded in lower hierarchical positions (Winter-Ebmer & Zweimuller, 1997); that women are held to higher promotion standards than men (Olson & Becker, 1983; Pekkarinen & Vartiainen, 2004); that women's promotion probabilities are more positively influenced by firm-specific job training (especially for younger and less tenured women), while men's probabilities are less affected (Melero, 2004); that women are less frequently in jobs that offer promotion opportunities than men, but when both genders are in jobs offering promotion opportunities, there is no significant gender difference in promotion (Groot & Maassen van den Brink, 1996); that women in the United States have roughly half the chance to be promoted to partner in the legal profession as men (Spurr, 1990); that women in a large Canadian corporation had a lesser chance of promotion than men after controlling for career-relevant factors (Cannings, 1988); and that no unexplained gender-specific differences in promotion existed in the US economics profession by the end of the 1980s (McDowell, Singell, & Ziliak, 2001).

10. Published papers employing these data include Abowd (1990), Bognanno (2001), and Belzil and Bognanno (2008).

11. Eriksson and Werwatz (2005) discuss the issues they faced in classifying jobs into levels for purposes of examining promotion in their multi-firm panel data set.

12. This approach is also common in empirical dynamic programming models with unobserved heterogeneity (Keane & Wolpin, 1997; Eckstein & Wolpin, 1999; Belzil & Hansen, 2002). It is largely influenced by the estimation method proposed by Heckman and Singer (1984).

13. One of the advantages of conditional likelihood techniques is the fact that statistical inference may be achieved without having to specify a distribution for the individual-specific effects, including the initial conditions. However, the conditional approach precludes the estimation of time invariant regressors (such as schooling), and does not allow one to recover the marginal effects.

14. The reader should remember that a smaller number for the level variable (L_{ijt}) implies a higher rank.

15. The ultimate choice regarding the number of types was based on the Hannan–Quinn information criterion (HQIC), which is

$$\text{HQIC} = \log L - \log(\log(N))\text{Þ} \tag{4}$$

where Þ refers to the number of parameters and L is the likelihood value.

16. In the applied literature, the number of support points allowed for when the heterogeneity term is a scalar rarely exceeds two. The decision here was based on the Hannan–Quinn information criterion, which penalizes the likelihood function for the number of parameters according to a penalty function defined as the log of the log sample size.

17. At the request of the journal, we provide the estimates of an OLS model for comparison. In the estimates below, the promotion variable is coded as the change in level, whether positive or negative, and not restricted to values of 0,1. Neither promotions nor demotions were restricted to being of one level. (Promotion = lag(level) − level, positive values reflect promotions and negative values reflect demotions.) As well, there were no recodings of promotion should, for instance, a promotion be followed by a demotion. All levels were included, even those greater than level 6. These points had almost no impact on the signs or significance of the parameter estimates relative to restricting promotion to a binary variable or excluding levels greater than six. In the estimation below, the parameter estimate is followed by the t-statistic in parentheses.

Promotion = −0.5400 + 0.0228 (20.1); Education − 0.0002 (0.8); Tenure − 0.0013 (2.7); Age − 0.0000 (1.5); Profits − 0.0000 (3.0); Sales − 0.0000 (4.3); Size + 0.0823 (46.0); Lag level − 0.2282 (9.0); Lag speed, $R^2 = 0.04$, $N = 88{,}426$.

The lags of level and speed are used because we wish to measure these prior to the promotion, not after they reflect it. These estimates are consistent with Table 1A, model 1 in regards to the effects of education, age, and level. Tenure is insignificant in both. The firm sales and size are not consistent. The coefficient on speed, computed as (12 − level)/(age − education − 4), is negative. We estimate in Table 3A a negative fast track effect that suggests for most people a high speed of past promotion hinders further promotion after conditioning on their ability. The negative sign on speed here is consistent with that finding.

18. It is possible to obtain level-specific parameters, but the results are more easily interpreted with a dichotomy between high-rank and low-rank levels.

19. To see this argument, consider estimating a model where the current promotion probability depends on the past two or three promotion outcomes. These parameter estimates would turn out to be negative and would imply that simulated promotion histories entail penalizing executives who have been promoted during the sampling period. Indeed, we have verified this assertion by estimating a dynamic promotion model where current promotions depend on up to three or four

past promotion outcomes. All parameters turn out to be negative, although those pertaining to order three and four were much weaker (very close to zero).

20. To verify this, we computed a marginal effect of promotion speed for an average individual. The marginal effect is equal to 0.0025.

21. As is the case in most nonlinear discrete dynamic models (such as in discrete choice panel data duration models), identification of a causal effect implicitly requires identification of the heterogeneity distribution. While a wide variety of nonparametric identification theorems exist in the literature, it is fair to say that virtually all of them require some form of separability between objects of interest. In our model, the separability between heterogeneity and the other portion of the choice determinant implies restrictions on the co-movement between heterogeneity and past promotions. For a full discussion see Honoré and Tamer (2006).

22. Consistent with a stronger promotion signal for the less educated, DeVaro and Waldman (2004) find empirical support for the promotion signaling implications that both the performance level required for promotion, and the wage increase upon promotion, are greater for the less educated.

23. Related papers in the literature include Sattinger (1975), Harris and Holmstrom (1982), Rosen (1982), Waldman (1984a, 1984b), Bernhardt (1995), and Farber and Gibbons (1996).

24. Meyer (1992) and Prendergast (1992) also derive fast tracks.

25. Gibbons and Waldman (2006) add explicit consideration of schooling to the model in a framework of symmetric learning. Worker ability is initially unknown and then gradually revealed through observations of output to all parties. Innate ability, θ_i, is increased by schooling. The interpretation is that schooling increases the speed at which workers learn from experience. More educated workers have a higher ability to learn on the job than less educated workers of equal labor market experience.

26. Other significant early papers with this information assumption include Ricart i Costa (1988) and Milgrom and Oster (1987).

ACKNOWLEDGMENT

We thank Jaap Abring, Francis Kramarz, Guy Laroque, Edward Lazear, Greg LeBlanc, Thierry Magnac, Bentley MacLeod, Gerard van den Berg, seminar participants at CREST, Tinbergen and HEC Lausanne, and three anonymous referees for useful comments. The support of CIRANO, a Temple University Research leave and a Marie Curie Fellowship for the Transfer of Knowledge are gratefully acknowledged by Bognanno. The usual disclaimer applies.

REFERENCES

Abowd, J. (1990). Does performance-based managerial compensation affect corporate performance? *Industrial and Labor Relations Review*, *43*(3), 52S–73S.

Abowd, J., Kramarz, F., & Margolies, D. (1999). High wage workers and high-wage firms. *Econometrica*, *67*(2), 251–333.

Ariga, K., Ohkusa, Y., & Brunello, G. (1999). Fast track: Is it in the genes? The promotion policy of a large Japanese firm. *Journal of Economic Behavior and Organization*, *38*, 385–402.

Baker, G., Gibbs, M., & Holmstrom, B. (1994a). The internal economics of the firm: Evidence from personnel data. *Quarterly Journal of Economics*, *CIX*, 921–955.

Baker, G., Gibbs, M., & Holmstrom, B. (1994b). The wage policy of a firm. *Quarterly Journal of Economics*, *CIX*, 881–919.

Belzil, C., & Bognanno, M. (2008). Promotions, demotions, halo effects, and the earnings dynamics of American executives. *Journal of Labor Economics*, *26*(2), 287–310.

Belzil, C., & Hansen, J. (2002). Unobserved ability and the return to schooling. *Econometrica*, *70*(5), 2075–2091.

Bernhardt, D. (1995). Strategic promotion and compensation. *The Review of Economic Studies*, *62*(2), 315–339.

Bognanno, M. (2001). Corporate tournaments. *Journal of Labor Economics*, *19*(2), 290–315.

Cannings, K. (1988). Managerial promotion: The effects of socialization, specialization, and gender. *Industrial and Labor Relations Review*, *42*(1), 77–88.

Chamberlain, G. (1984). Panel data. In: Z. Griliches & M. D. Intriligator (Eds), *Handbook of econometrics* (Vol. II, Chapter 22). New York, NY: Elsevier.

Chiappori, P.-A., Salanie, B., & Valentin, J. (1999). Early starters versus late beginners. *Journal of Political Economy*, *107*(4), 731–760.

DeVaro, J. (2006). Internal promotion competitions. *Rand Journal of Economics*, *37*(3), 521–542.

DeVaro, J., & Brookshire, D. (2007). Promotions and incentives in non-profit and for-profit organizations. *Industrial and Labor Relations Review*, *60*(3), 311–339.

DeVaro, J., & Waldman, M. (2004). The signaling role of promotions: Further theory and empirical evidence. Mimeo, Cornell University, Ithaca, NY (December 2004).

Dohmen, T. J., Kriechel, B., & Pfann, G. A. (2004). Monkey bars and ladders: The importance of lateral and vertical job mobility in internal labor market careers. *Journal of Population Economics*, *17*(2), 193–228.

Eckstein, Z., & Wolpin, K. (1999). Why youth drop out of high school: The impact of preferences, opportunities and abilities. *Econometrica*, *67*(6), 1295–1339.

Eriksson, T., & Werwatz, A. (2005). The prevalence of internal labour markets – New evidence from panel data. *International Journal of Economics Research*, *2*(2), 105–124.

Farber, H. S., & Gibbons, R. (1996). Learning and wage dynamics. *The Quarterly Journal of Economics*, *111*(4), 1007–1047.

Forbes, J. B., & Piercy, J. E. (1991). *Corporate mobility and paths to the top*. New York: Quorum Books.

Gibbons, R., & Waldman, M. (1999). A theory of wage and promotion dynamics inside firms. *The Quarterly Journal of Economics*, *114*(4), 1321–1358.

Gibbons, R., & Waldman, M. (2006). Enriching a theory of wage and promotion dynamics inside firms. *Journal of Labor Economics*, *24*(1), 59–107.

Gibbs, M., & Hendricks, W. E. (2004). Do formal salary systems really matter? *Industrial and Labor Relations Review*, *58*(1), 71–93.

Groot, W., & Maassen van den Brink, H. (1996). Glass ceilings or dead-ends: Job promotion of men and women compared. *Economic Letters*, *53*, 221–226.

Harris, M., & Holmstrom, B. (1982). A theory of wage dynamics. *The Review of Economic Studies*, *49*(3), 315–333.

Heckman, J. (1981). Statistical models for discrete panel data. In: C. Manski & D. McFadden (Eds), *Structural analysis of discrete data with economic application*. Cambridge, MA: MIT Press.

Heckman, J., & Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, *52*(2), 271–320.

Honoré, B., & Tamer, E. (2006). Bounds on parameters in panel dynamic discrete choice models. *Econometrica*, *74*(3), 611–629.

Howard, A., & Bray, D. W. (1988). *Managerial lives in transition: Advancing age and changing times*. New York: Guilford.

Keane, M. P., & Wolpin, K. I. (1997). The career decisions of young men. *Journal of Political Economy*, *105*(3), 473–522.

Lazear, E. (1992). The job as a concept. In: W. J. Bruns, Jr. (Ed.), *Performance measurement, evaluation, and incentives*. Boston, MA: Harvard Business School Press.

Magnac, T. (2000). Subsidized training and youth employment: Distinguishing unobserved heterogeneity from state dependence in labour market histories. *The Economic Journal*, *110*(466), 805–837.

McCue, K. (1996). Promotions and wage growth. *Journal of Labor Economics*, *14*(2), 175–209.

McDowell, J. M., Sengell, L. D., Jr., & Ziliak, J. (2001). Gender and promotion in the economics profession. *Industrial and Labor Relations Review*, *54*(2), 224–244.

Melero, E. (2004). *Evidence on training and career paths: Human capital, information and incentives*. IZA DP no. 1377 (November 2004). Bonn, Germany.

Meyer, M. (1992). Biased contests and moral hazard: Implications for career profiles. *Annales d' Economie et de Statistique*, *25*, 165–187.

Milgrom, P., & Oster, S. (1987). Job discrimination, market forces, and the invisibility hypothesis. *The Quarterly Journal of Economics*, *102*(3), 453–476.

Olson, C. A., & Becker, B. E. (1983). Sex discrimination in the promotion process. *Industrial and Labor Relations Review*, *36*(4), 624–641.

Pekkarinen, T., & Vartiainen, J. (2004). *Gender differences in job assignment and promotion on a complexity ladder of jobs*. IZA DPN no. 1184 (June 2004). Bonn, Germany.

Pergamit, M., & Veum, J. (1999). What is a promotion? *Industrial and Labor Relations Review*, *52*(4), 581–601.

Prendergast, C. (1992). Career development and specific human capital collection. *Journal of the Japanese and International Economics*, *6*, 207–227.

Ricart i Costa, J. E. (1988). Managerial task assignment and promotions. *Econometrica*, *56*(2), 449–466.

Rosen, S. (1982). Authority, control, and the distribution of earnings. *The Bell Journal of Economics*, *13*(2), 311–323.

Rosenbaum, J. E. (1979). Tournament mobility: Career patterns in a corporation. *Administrative Science Quarterly*, *24*, 220–241.

Rosenbaum, J. E. (1984). *Career mobility in a corporate hierarchy*. Orlando: Academic Press, Inc.

Sattinger, M. (1975). Comparative advantage and the distributions of earnings and abilities. *Econometrica*, *43*(3), 455–468.

Seltzer, A., & Merrett, D. (2000). Personnel policies at the Union Bank of Australia: Evidence from the 1888–1900 entry cohorts. *Journal of Labor Economics*, *18*(4), 573–613.

Spurr, S. (1990). Sex discrimination in the legal profession: A study of promotion. *Industrial and Labor Relations Review*, *43*(4), 406–417.
Treble, J., van Gameren, E., Bridges, S., & Barnby, T. (2001). The internal economics of the firm: Further evidence from personnel data. *Labour Economics*, *8*(5), 531–552.
Tuckel, P., & Siegel, K. (1983). The myth of the migrant manager. *Business Horizons*, *26*(1), 64–70.
Useem, M., & Karabel, J. (1986). Pathways to top corporate management. *American Sociological Review*, *51*, 184–200.
Vroom, V. F., & MacCrimmon, K. R. (1968). Toward a stochastic model of managerial careers. *Administrative Science Quarterly*, *13*, 26–46.
Waldman, M. (1984a). Job assignments, signalling, and efficiency. *The Rand Journal of Economics*, *15*(2), 255–267.
Waldman, M. (1984b). Worker allocation, hierarchies and the wage distribution. *The Review of Economic Studies*, *51*(1), 95–109.
Winter-Ebmer, R., & Zweimuller, J. (1997). Unequal assignment and unequal promotion in job ladders. *Journal of Labor Economics*, *15*(1), 43–71.
Wise, D. A. (1975). Personal attributes, job performance and probability of promotion. *Econometrica*, *43*(5–6), 913–931.
Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, *20*(1), 39–54.

DATA APPENDIX

Table A1. Summary Statistics for Executives in Levels 1–6.

Variables	Mean	SD
Education (years)	16.42	1.85
Age	46.21	8.66
Tenure	13.24	10.31
Newcomers in current year	0.036	0.19
Promotions in 2nd year	0.11	0.31
Initial speed of promotion	0.37	0.21

Table A2. Summary Statistics by Individual Observation Number.

Individual Observation	Level		Fraction Promoted
	Mean	SD	Mean
1st	4.18	1.15	–
2nd	4.18	1.19	0.11
3rd	4.09	1.22	0.09
4th	3.99	1.22	0.08
5th	3.92	1.25	0.06
6th	3.82	1.27	0.09
7th	3.66	1.23	0.08
8th	3.48	1.16	0.05

Table A3. Promotion Incidence by Level in the First Two Years of Data for Each Individual.

Level	Number of Individuals	Fraction Promoted	Average Tenure	Average Age	Base and Bonus (1980$)
1	316	–	21.1	55.8	400,026
2	1,951	0.011	15.1	50.4	158,294
3	6,473	0.043	12.9	47.4	91,356
4	10,111	0.081	12.8	45.9	67,559
5	8,207	0.150	13.1	45.1	57,300
6	4,384	0.216	13.6	44.6	49,181
7	1,488	0.265	14.7	44.6	43,250
8	380	0.350	14.9	43.8	37,906
9	161	0.330	15.2	43.8	30,374
10	54	0.370	15.2	41.6	25,363
11	12	0.417	18.8	44.1	49,703
All	33,537	0.117	13.3	46.1	74,148

THEORY APPENDIX: FAST TRACKS

A theoretical motivation for examining fast tracks comes from the job assignment literature. The model that we present is contained in Gibbons and Waldman (1999) and draws upon several earlier papers.[23] We focus on the theoretical implications from this paper in the case of full information and Bernhardt (1995) in the case of asymmetric learning because they both model three job levels and imply promotion fast tracks.[24] The basic model is

Table A4. Firm Level Statistics by Year.

Year	Mean Executive Observations Per Firm		Mean Levels Reported on Per Firm		Profits (Million 1980$)		Sales (Million 1980$)		Size (Employees)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1981	76	48	6.3	1.6	152	422	2,989	5,554	30,625	44,666
1982	81	55	6.2	1.7	163	485	3,035	7,118	31,525	69,013
1983	80	57	6.1	1.7	104	266	2,672	5,759	27,414	38,684
1984	80	68	6.1	1.5	98	272	2,359	4,770	25,985	36,156
1985	80	60	6.1	1.5	120	308	2,562	5,133	28,326	36,380
1986	81	57	6.1	1.6	114	280	2,740	5,031	30,619	39,648
1987	81	62	5.9	1.6	132	296	2,804	4,502	31,075	44,744
1988	78	50	5.8	1.5	121	335	2,767	4,738	29,806	45,103

Table A5. Promotion Incidence in the First Two Years of Data for Individuals Younger and Older than Average in their Firm and Level.

	Number of Individuals	Mean	SD
All levels			
Younger	17,933	0.120	0.325
Older	14,501	0.114	0.317
Levels 2, 3			
Younger	4,280	0.032	0.175
Older	3,618	0.033	0.178
Levels 4, 5, 6			
Younger	12,574	0.134	0.340
Older	10,007	0.129	0.335

followed by the implications derived in the literature with alternative assumptions made in regards to information.

Identical firms in a competitive market with free entry, producing with only labor, assign workers to three exogenously determined jobs. Output in each job consists of two components, one that is independent of the worker in the job and one that depends on the effective ability of the worker. The value of the labor market experience gained by the worker is the same across jobs. Effective ability depends on the innate ability of the worker and on labor market experience. The parameters determining output in the three jobs are set so as to differentially value effective ability such that workers, as

they gain experience, progress through the jobs sequentially. Because workers differ in innate ability, they grow in effective ability with labor market experience at different rates and therefore have different speeds of promotion.

The labor market experience of worker i in period t is denoted x_{it}. The worker's innate ability is represented by θ_i and effective ability by

$$\eta_{it} = \theta_i f(x_{it}) \tag{A.1}$$

where $f' > 0$, $f'' \leqslant 0$ and $f(0) = 0$. The output of worker i at time t in job j ($j = 1, 2, 3$) is

$$y_{ijt} = d_j + c_j(\eta_{it} + \varepsilon_{ijt}) \tag{A.2}$$

where the constants d_j and c_j are such that $d_1 > d_2 > d_3 > 0$ and $0 < c_1 < c_2 < c_3$ and $\varepsilon_{ijt} \sim N(0, \sigma^2)$ is an error term. We define η' and η'' to indicate the effective abilities at which a worker's expected output is equal between jobs 1 and 2 and jobs 3 and 4 respectively. Hence, $d_1 + c_1\eta' = d_2 + c_2\eta'$ and $d_2 + c_2\eta'' = d_3 + c_3\eta''$.

Given this structure, we can actually compute the theoretical probability of promotion at time t for each level. For a given population distribution function of innate ability $G(\theta)$ combined with Eq. (A.1), we may obtain the nonstationary distribution function of η_{it}, which we denote by $G_{\eta t}(.)$. Given $G_{\eta t}(.)$, the population density of individuals promoted to level 2 at time t is

$$\Pr(\eta_{it} \geqslant \eta' \mid \eta_{it-1} < \eta') = \frac{1 - G_{nt}(\eta')}{G_{\eta t-1}(\eta')} \tag{A.3}$$

while the population density of individuals promoted to level 3 at time t is

$$\Pr(\eta_{it} \geqslant \eta'' \mid \eta' < \eta_{it-1} < \eta'') = \frac{1 - G_{nt}(\eta'')}{G_{\eta t-1}(\eta'') - G_{\eta t-1}(\eta')} \tag{A.4}$$

Without any further parametric assumptions, it is not possible to say how the promotion probability evolves with level. In other words, the relative difficulty of subsequent promotions cannot be determined in standard job assignment models.

Full Information

In a world of full information (θ_i is public knowledge), the assignment of workers maximizes their expected output. Assignment is to job 1 when

effective ability, η_{it}, is below η' and to job 2 when η_{it} surpasses η' but remains below η'' and to job 3 when η_{it} surpasses η''. Entering workers ($x_{it} = 0$) are always assigned to job 1 since $f(0) = 0$. With full information and competitive markets, wages are set such that $w_{ijt} = d_j + c_j(\eta_{it})$.

The full information model has implications for both the effect of schooling on promotion, when schooling is a component of θ_i, and fast tracks. The parametric assumption that effective ability, η_{it}, is multiplicative in θ_i and labor market experience, $f(x_{it})$, implies that schooling increases the growth rate of effective ability and the rate of promotion.[25] More generally, it implies that workers endowed with higher θ_i attain the threshold levels of effective ability required for promotion to the next level, η' and η'', faster than those with less innate ability. Achieving one promotion relatively quickly will be correlated with achieving the next promotion relatively quickly. We label the fast tracks in the full information model *noncausal*. This terminology indicates that a rapid promotion doesn't cause a subsequent rapid promotion. Rather, high innate ability is the underlying source of the high speed of promotion. In econometric terms, this means that fast tracks in the presence of full information are solely explained by individual heterogeneity.

As more able people are promoted more quickly, the time it takes to achieve promotion is indicative of innate ability. Controlling for level, workers with less labor market experience will be of greater innate ability than those who required more time to achieve the level. We expect executives who are younger than average in their level at their specific firm to be of higher innate ability and therefore more promotable.

Asymmetric Learning

Waldman (1984a) was the first to model outside firms learning about worker ability through the signal provided by job assignment.[26] Suppose that the current firm is able to perfectly observe worker ability after the initial job assignment but that outside firms can only observe the worker's current and past job assignment, education, and wages. In this framework, wages are no longer equal to the expected value of the worker's actual production in the job assigned to maximize output. Instead, wages are equal to the value of the worker as perceived by outside labor market based only on public information. As outside firms infer higher ability to more educated and rapidly promoted workers, such workers must be paid more by their current firm to avoid being bid away.

Workers from high-ability groups will receive more promotions and at a faster pace than equally (or more) able workers from low-ability groups. When workers from high-ability groups are promoted, the public perception of their ability rises less than when someone from a low-ability group is promoted. Since an increase in perceived ability must be met with an increased wage, firms have a bias in favor of promoting those from high-ability groups. Firms are able to exploit high-ability workers from low-ability groups through underpayment and delayed promotion when their ability has not been signaled to outside firms.

Accordingly, promoting educated workers changes the perception of ability less than promoting uneducated workers. Bernhardt (1995) shows that firms will have a bias in favor of promoting educated workers over equally able (or more able) uneducated workers because the wage revision is smaller when promoting those from a more able population. Because the signaling role of promotion is stronger for workers with lower levels of education, those promoted with less education must be exceptionally able.

Bernhardt motivates fast tracks under asymmetric learning with reasoning similar to the discussion above. If a worker is promoted faster than other workers to the next level, the worker is signaled to be more able. Even if the slower workers reach the same level and begin to outperform, the firm has an incentive to promote the faster worker first to the subsequent level. Rapid early promotion signals high ability (and this signal is stronger for less educated workers) and makes further rapid promotions more likely. Fast tracks resulting for this reason we label *causal* to indicate that early promotions have their own inherent effect on the pace of subsequent promotions that is independent of ability. Noncausal fast tracks also result in this framework as, ceteris paribus, more able workers maintain a higher speed of promotion. In econometric terms, fast tracks in the presence of asymmetric information are explained by individual heterogeneity and past promotion histories.

This discussion of fast tracks being rooted in rapid early promotion suggests the importance of early career results in providing signals to the outside labor market. With promotion, the importance of originating from a high or low-ability group becomes less relevant as outside firms draw conclusions based on the employment history. In the context of a model with two job levels, Bernhardt states that once an able worker from a low-ability group has been promoted, the worker can no longer be exploited because the worker's ability has been revealed to outside firms. Observable human capital variables, such as years of education, may be expected to have a diminishing impact on promotion probabilities at higher levels if their role is limited to providing a signal of ability and not one of increasing the ability to learn on the job.

SELF-SELECTION MODELS FOR PUBLIC AND PRIVATE SECTOR JOB SATISFACTION

Simon Luechinger, Alois Stutzer and
Rainer Winkelmann

ABSTRACT

We discuss a class of copula-based ordered probit models with endogenous switching. Such models can be useful for the analysis of self-selection in subjective well-being equations in general, and job satisfaction in particular, where assignment of regressors may be endogenous rather than random, resulting from individual maximization of well-being. In an application to public and private sector job satisfaction, and using data on male workers from the German Socio-Economic Panel for 2004, and using two alternative copula functions for dependence, we find consistent evidence for endogenous sector selection.

1. INTRODUCTION

The distinction between public and private sector employment conditions has generated a sizeable literature in empirical labor economics, the largest part of which has studied the wage structure in the two sectors.

Jobs, Training and Worker Well-Being
Research in Labor Economics, Volume 30, 233–251
Copyright © 2010 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 0147-9121/doi:10.1108/S0147-9121(2010)0000030010

A key concern for any study in this area is the potential non-random selection of workers into sectors which renders the comparison of outcomes for public sector workers and private sector workers uninformative for the causal effect of sector affiliation on wages. The resulting endogeneity problem has been addressed in one of two ways, either by following workers over time and including fixed individual effects (e.g., Pederson, Schmidt-Sorensen, Smith, & Westergard-Nielsen, 1990), or by specifying a switching regression model for cross-sectional data (e.g., van der Gaag & Vijverberg, 1988; Zweimüller & Winter-Ebmer, 1994; Dustmann & Van Soest, 1998).

Both strategies have been borrowed in more recent studies that consider job satisfaction, rather than wages, as the outcome variable of interest. For example, Heywood, Siebert, and Wei (2002) use panel data from the British Household Panel Study and conclude that public sector workers are "positively selected," meaning that the public sector attracts workers who are more easily satisfied anyway. If the sorting of workers is driven by idiosyncratic gains from being in one sector rather than the other, however, such fixed effects models are inappropriate. The switching regression approach allows for selection effects driven by relative gains in job satisfaction. This is a likely scenario if workers are heterogeneous in their preferences for job attributes offered in the two sectors.

Nevertheless, previous implementations for job satisfaction have been rare. This may be due to the fact that standard switching regression models are tailored to continuous-dependent variables, whereas job satisfaction is a discrete and ordered outcome. Asiedu and Folmer (2007) use a two-step approach where regressors in an ordered probit model for job satisfaction in each sector are augmented by a predicted inverse Mills ratio. McCausland, Pouliakas, and Theodossiou (2005) disregard the discreteness of the job satisfaction response and use a standard linear model.

The alternative followed in this chapter is to specify a linear switching regression for *latent* continuous outcomes, and specify a threshold mechanism that translates the latent model into corresponding discrete ordered response probabilities. If the stochastic errors in the latent model are jointly normal distributed, a multivariate ordered probit model results (e.g., Greene & Hensher, 2008; Munkin & Trivedi, 2008; the frequently used bivariate probit model is a special case). We show, how alternative dependence structures can be modeled in a copula framework.

The rest of the chapter is organized as follows. The next section develops the essential elements of a switching-regression model for job satisfaction. Section 3 introduces copulas as a natural characterization of dependence

in such a switching regression model. The general likelihood function is derived, and three-specific cases are considered: independence copula, normal copula, and Frank's copula. Section 4 applies the copula method to job satisfaction of public and private sector workers. Tests show that the Frank copula dominates the other models in this application. Falsely ignoring self-selection means that the effect of sector allocation on job satisfaction is underestimated. Section 5 concludes the chapter.

2. MODELING SELF-SELECTION IN JOB SATISFACTION

When studying subjective well-being and its domains, including job satisfaction, self-selection arises naturally, since one can expect rational individuals to choose their life circumstances with a view toward maximizing well-being. This has to be recognized when attempting to estimate the effect of a choice variable on satisfaction. In this chapter, we consider the choice between public and private sector employment, and its effect on job satisfaction.

Let $U_i(1)$ be the job satisfaction of a person working in sector 1, the public sector, while $U_i(0)$ is the job satisfaction of the same worker while working in sector 0, the private sector. By construction, one of the two outcomes is unobservable. For public sector workers, we can observe $U_i(1)$ but not $U_i(0)$, and vice versa for private sector workers. Hence, the public–private sector job satisfaction differential for worker i, $U_i(1) - U_i(0)$, is unidentified. In principle, we can attempt to identify population averages, such as $\mathrm{E}[U_i(1) - U_i(0)]$ (the average treatment effect).

Assume that people choose the sector where they expect to be most satisfied, and their expectations are fulfilled. The realized sector is denoted by $s \in \{0, 1\}$, where $s_i = 0$ means that worker i works in the private sector, and $s_j = 1$ means that worker j works in the public sector. Under the above assumption, $s_i = 0$ if and only if $U_i(1) < U_i(0)$ and $s_j = 1$ if and only if $U_j(1) > U_j(0)$. As a consequence, we can identify $\mathrm{E}[U_i(1)|U_i(1) > U_i(0)]$, but, without further assumptions, not $\mathrm{E}[U_i(1)]$. Similarly, we can identify $\mathrm{E}[U_i(0)|U_i(1) < U_i(0)]$, but not $\mathrm{E}[U_i(0)]$. Ignoring this issue leads to selection bias. For example, the coefficient of a sector 1 dummy variable in a regression model will not typically estimate the average treatment effect as defined above.

2.1. A Switching Regression Model of Job Satisfaction

One possible set of assumptions that enable estimation of the effect of sector on job satisfaction, while controlling for a number of explanatory variables, is offered by the standard switching regression model that can be adjusted in order to account for the discrete and ordered response, job satisfaction. Let

$$y_0^* = x'\beta_0 + \varepsilon_0 \tag{1}$$

be the latent job satisfaction index if $s = 0$, and

$$y_1^* = x'\beta_1 + \varepsilon_1 \tag{2}$$

be the latent job satisfaction index if $s = 1$. x is a vector of explanatory variables that is the same in both equations, and β_0, β_1 are conformable sector-specific parameter vectors. We do not impose that $\beta_0 = \beta_1$, that is, the regression coefficients may be sector-specific. Workers are observed either in sector $s = 1$ or in sector $s = 0$, but never in both at the same point in time. It is unreasonable to assume that workers select themselves randomly into the sectors. Rather, it is likely that there is self-selection based on idiosyncratic gains to job satisfaction due to preference heterogeneity. For example, workers who gain most from being in the public sector are actually the ones choosing $s = 1$ with highest probability. Selection is captured by a third latent equation,

$$s^* = z'\gamma + v \tag{3}$$

and

$$s = \begin{cases} 1 & \text{if } s^* \geq 0 \\ 0 & \text{if else} \end{cases} \tag{4}$$

Usually, in this kind of model, z includes a number of instruments in addition to x. The reason x should be a subset of z is that x affects sector-specific job satisfaction, which is likely to be a factor in determining a person's sectoral choice. Exclusion restrictions are required in order to identify the model in other ways rather than through functional form assumptions on the error term only.

The observation mechanism is completed by accounting for the discrete and ordinal scale of *observed* job satisfaction. In particular, we follow

standard practice and assume a threshold observation mechanism, whereby

$$y_s = \sum_{j=0}^{J} \mathbb{1}(y_s^* > \kappa_{s,j}), \quad s = 0, 1$$

and $\kappa_{s,0} = -\infty < \kappa_{s,1} < \ldots < \kappa_{s,J} = \infty$ partition the real line (i.e., $y_s = j$ if and only if $\kappa_{s,j-1} < y_s^* \leq \kappa_{s,j}, j = 1, 2, \ldots, J$). This is not a standard ordered response model since y_s is only partially observed. Observed job satisfaction is obtained as

$$y = y_0^{1-s} y_1^s$$

Based on the latent model structure, the probabilities of observed private and public sector job satisfaction can be written as

$$\begin{aligned} P(y_0 = j, s = 0|x, z) &= P(\kappa_{0,j-1} - x'\beta_0 < \varepsilon_0 \leq \kappa_{0,j} - x'\beta_0, v \leq -z'\gamma) \\ &= P(\varepsilon_0 < \kappa_{0,j} - x'\beta_0, v \leq -z'\gamma) \\ &\quad - P(\varepsilon_0 < \kappa_{0,j-1} - x'\beta_0, v \leq -z'\gamma) \end{aligned} \tag{5}$$

and

$$\begin{aligned} P(y_1 = j, s = 1|x, z) &= P(\kappa_{1,j-1} - x'\beta_1 < \varepsilon_1 \leq \kappa_{1,j} - x'\beta_1, v > -z'\gamma) \\ &= P(\varepsilon_1 < \kappa_{1,j} - x'\beta_1) - P(\varepsilon_1 < \kappa_{1,j-1} - x'\beta_1) \\ &\quad - P(\varepsilon_1 < \kappa_{1,j} - x'\beta_1, v \leq -z'\gamma) \\ &\quad + P(\varepsilon_1 < \kappa_{1,j-1} - x'\beta_1, v \leq -z'\gamma) \end{aligned} \tag{6}$$

In this model, the absence of self-selection is equivalent to statistical independence of v and ε_0 and ε_1, respectively. With independence, the joint probabilities can be factored into their marginals, and one obtains univariate ordered and binary response models. The nature of self-selection, if present, correspondingly hinges on the joint distributions $f(v, \varepsilon_0)$ and $f(v, \varepsilon_1)$. For example, if v and ε_0, and v and ε_1, are bivariate normally distributed, with correlations ρ_0 and ρ_1, respectively, the model has a multivariate ordered probit structure (where the correlation between ε_0 and ε_1 is unidentified). The marginal models for sector-specific job satisfaction are ordered probits, and the selection model is a binary probit.

But even if one wants to keep probit marginals for all three equations, the two joint distributions do not need to be bivariate normal. We suggest to combine the outlined switching regression model with a copula approach for generating joint distribution functions for given marginals. In this way,

we can potentially specify many ordered probit models with endogenous switching in a unified framework.

Copulas have been used in econometrics before but, to the best of our knowledge, so far not in the present context of ordered responses. A brief history and overview of the technique is given in the next section, before we return to the specific implementation of a model for job satisfaction under self-selection.

3. MODELING SELECTION USING COPULAS

Copulas offer a particular representation of arbitrary joint distribution functions, with the key property being that the specification of the marginal distributions and the dependence structure is "uncoupled." The earliest copula use in econometrics was by Lee (1983) who suggested, in the context of the sample selection model, to use a bivariate normal copula (more on this below) for generating dependence between two continuous random variables, one with normal marginal (the continuous outcome variable) and one with logistic distribution (the error in the latent selection equation). The first econometric applications to discrete outcomes were provided by van Ophem (1999, 2000) who used a bivariate normal copula to generate joint distributions for two random variables with Poisson/Poisson and Poisson/normal marginals, respectively.

The systematic consideration of non-normal copulas started with Smith (2003) who specified eight different copulas for normal/normal and normal/gamma marginals. Further contributions in this area include Smith (2005) who used five different copulas in a switching regression model for continuous outcomes, and Zimmer and Trivedi (2006) who used the Frank copula for negative binomial/normal marginals. An introduction to the copula method for empirical economists is provided by Trivedi and Zimmer (2007), see also Nelson (2006).

In statistics, a two-copula is a bivariate joint distribution function defined on the two-dimensional unit cube [0,1] such that both marginal distributions are uniform on the interval [0,1]. For example, the normal, or Gaussian, family of copulas, for $n = 2$, is

$$P(U \leq u, V \leq v) = C(u, v) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v); \rho) \tag{7}$$

where Φ and Φ_2 are the uni- and bivariate cdf of the standard normal distribution, and $-1 \leq \rho \leq 1$ is the coefficient of correlation. Another

example is the Frank family of copulas

$$C(u,v) = -\theta^{-1}\log\left\{1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{(e^{-\theta}-1)}\right\} \quad -\infty < \theta < \infty \tag{8}$$

A comprehensive summary of copulas is provided by Nelson (2006). The marginal distributions implied by bivariate copulas are

$$F(u) = P(U \le u, V \le 1) = C(u,1)$$

and

$$F(v) = P(U \le 1, V \le v) = C(1,v)$$

respectively. It is easy to verify that all three copulas have the key property that their marginal distributions are uniform, as $C(u,1) = u$ and $C(1,v) = v$.

The significance of copulas lies in the fact that by way of transformation, any joint distribution function can be expressed as a copula applied to the marginal distributions. This result is due to Sklar (1959). Sklar's theorem states that given a joint distribution function $F(y_1,\ldots,y_k)$, and respective marginal distribution functions, there exists a copula C such that the copula binds the margins to give the joint distribution.

For the bivariate case, Sklar's theorem can be stated as follows. For any bivariate distribution function $F(y_1,y_2)$, let $F_1(y_1) = F(y_1,\infty)$ and $F_2(y_2) = F(\infty,y_2)$ be the univariate marginal probability distribution functions. Then there exists a copula C such that

$$F(y_1,y_2) = C(F_1(y_1),F_2(y_2))$$

Moreover, if the marginal distributions are continuous, the copula function C is unique. We see that the copula is now expressed as a function of cdfs. But cdfs are uniformly distributed over the interval [0, 1]. Since the marginal distributions of a copula are uniform, it follows that the marginal distributions of $y_1 = F_1^{-1}(u)$ and $y_2 = F_2^{-1}(v)$ are F_1 and F_2, as stated.

The practical significance of copula functions in empirical modeling stems from the fact that they can be used to build new multivariate models for given univariate marginal component cdfs. If the bivariate cdf $F(y_1,y_2)$ is unknown, but the univariate marginal cdfs are of known form, then one can choose a copula function and thereby generate an approximation to the unknown joint distribution function. The key is that this copula function introduces dependence, captured by additional parameter(s), between the two random variables (unless the independence copula $C(u,v) = uv$ is chosen). The degree and type of dependence depends on the choice of copula

family as well as the parameters. For our purposes, it is essential that the copula allows for positive *and* negative correlation, since we do not want to restrict the selection pattern a priori: we want to learn from the data whether workers observed in sector 1 are more, less, or equally satisfied in comparison to a randomly selected worker in that sector, *ceteris paribus*, that is, for a given set of explanatory variables.

We consider three copula functions in the following application, the normal copula, the Frank copula, and the independence copula $C(u, v) = uv$. In the normal case, $-1 \leq \rho \leq 1$, with -1 signifying perfect negative correlation, 0 signifying independence, and $+1$ signifying perfect positive correlation. Since copulas in general do not impose linear dependence structures, correlation measures have only limited information value when moving away from the normal copula. There are a number of other indicators of a copula's ability to generate dependence (see Trivedi & Zimmer, 2007, for a detailed discussion). One is the question whether it can reach the Fréchet upper and lower bounds. The Fréchet upper bound for any bivariate distribution is given by $F_{\mathrm{u}}(y_1, y_2) = \min[F_1(y_1), F_2(y_2)]$, where F_1 and F_2 are the marginal cdfs. $F(y_1, y_2) = F_{\mathrm{u}}$ requires F to be the most positive-dependent bivariate distribution in any possible sense. The lower bound is given by $F_{\mathrm{l}}(y_1, y_2) = \max[0, F_1(y_1) + F_2(y_2) - 1]$, representing greatest possible negative dependence. Both normal and Frank copula can reach F_{l} and F_{u}, and thus span the full range of dependence. For the Frank copula, the dependence parameter may assume any real value. Values of $-\infty$, 0, and ∞ correspond to the Fréchet lower bound, independence, and the Fréchet upper bound, respectively. Like the normal copula, the Frank copula is symmetric in both tails.

3.1. Implementation for Ordered Response Models

For any given copula, the two required joint probabilities, $P(y_0 = j, s = 0|x, z)$ and $P(y_1 = j, s = 1|x, z)$ in Eqs. (5) and (6) are fully determined up to the unknown parameters. The assumption of ordered probit and probit marginals requires that $v \sim \text{Normal}(0, 1)$, $\varepsilon_1 \sim \text{Normal}(0, 1)$, $\varepsilon_0 \sim \text{Normal}(0, 1)$, where the variances are normalized to unity for identification. Thus,

$$\begin{aligned} P(y_0 = j, s = 0|x, z) = {} & C(\Phi(\kappa_{0,j} - x'\beta_0), \Phi(-z'\gamma), \theta_0) \\ & - C(\Phi(\kappa_{0,j-1} - x'\beta_0), \Phi(-z'\gamma), \theta_0) \end{aligned} \tag{9}$$

and

$$P(y_1 = j, s = 1|x, z) = C(\Phi(\kappa_{1,j} - x'\beta_1), 1, \theta_1) - C(\Phi(\kappa_{1,j-1} - x'\beta_1), 1, \theta_1)$$

$$-C(\Phi(\kappa_{1,j} - x'\beta_1), \Phi(-z'\gamma), \theta_1) + C(\Phi(\kappa_{1,j-1} - x'\beta_1), \Phi(-z'\gamma), \theta_1) \quad (10)$$

where $C(u, v)$ is either the normal copula (Eq. 7), Frank's copula (Eq. 8), or the independence copula. The parameters of the model, $\xi = (\kappa_0, \kappa_1, \beta_0, \beta_1, \gamma, \theta_0, \theta_1)'$, can be estimated by maximum likelihood, or quasi-maximum likelihood. Given an independent sample of observation tuples (y_i, s_i, x_i, z_i), the likelihood function is simply

$$L(\xi; y, s, x, z) = \prod_{i=1}^{n} P(y_s, s|x, z) \quad (11)$$

In our application, the log-likelihood function was maximized using the MAXLIK routine in GAUSS with numerical first and second derivatives. No convergence problems were encountered. Under the assumptions of the model, the maximum-likelihood estimator has the desirable large sample properties. If the model is misspecified, it is a quasi-likelihood estimator in the sense of White (1982), that is the best approximation (in a Kullback–Leibler sense) to the true model.

The normal and Frank specifications are non-nested and information criteria can be used to select among competing models. Alternatively, Vuong (1989) provides a framework for formal testing. Since the two models are overlapping, both including the independence copula as a special case, the two-step procedure should be applied.

The estimated ordered probit coefficients have the usual interpretation related to such models (see, for instance, Boes & Winkelmann, 2006). In particular, they can be used to compute marginal effects for a randomly selected worker in the two sectors, net of selection bias. A comparison of the outcome distribution of a randomly selected worker in the two sectors provides an estimate of the average treatment effect.

The dependence parameters θ_s inform about the direction of the selection bias. The null hypothesis of no self-selection implies that $\theta_s = 0$, a hypothesis that can be tested directly. If rejected, an interesting quantification of the selection effects can be obtained by comparing the outcome distribution of self-selected workers, for instance $p_{01} = P(y_0 = j|s = 1, x, z)$, with the *counterfactual* predicted distribution $p_{00} = P(y_0 = j|s = 0, x, z)$ of a worker who chooses state 1 but is (hypothetically) allocated to sector 0. For instance, positive selection is defined as a situation where p_{01} lies to the

right of p_{00}, in the sense that the probability of reporting high levels of job satisfaction in sector 1 is higher for workers who actually chose that sector, relative to others.

4. JOB SATISFACTION OF PUBLIC AND PRIVATE SECTOR WORKERS IN GERMANY

In this section, the copula methodology is applied to a model of sectoral job satisfaction in West Germany. We distinguish between two sectors, the private sector and the public (or government) sector. The question of empirical interest in this application is whether sector-specific job satisfaction and sector choice are jointly determined. If so, public (and private) sector workers are not representative of the entire population of workers. As a consequence, estimating a model of public sector job satisfaction using public sector workers, or of private sector job satisfaction using private sector workers, does not recover the underlying population relationships. For instance, such sub-sample estimates would misrepresent the job satisfaction difference between the two sectors for an average worker. Specifically, we suspect selection based on comparative gain, whereby public sector workers are those who gain most from that type of work environment, whereas private sector workers are those whose preferences and values are better matched in private sector jobs.

The selection effects we are interested in are conditional on other observed determinants. The general latent variable model was formulated in Eqs. (1) and (2) as

$$y_s^* = x'\beta_s + \varepsilon_s \quad s = 0, 1$$

where $s = \mathbb{1}(z'\gamma + v > 0)$. Moreover, y_s^* is the latent job satisfaction index in the private ($s = 0$) and public ($s = 1$) sector, respectively, and x is a vector of explanatory variables that affects job satisfaction. We estimate all models with two different sets of regressors. In a first model, we only include worker-specific covariates, similar to those found in related papers on the topic of job satisfaction (e.g., Clark, 1997). In a second model, we add to those worker-specific covariates a set of job-specific attributes, such as working hours, wages, and firm size. The two models answer different questions that both are of independent interest. The second model determines the effect of working in the public sector on satisfaction *conditional* on certain job attributes, that is, for a job in a similar sized firm, paying the same wage

and requiring the same working hours. In the first model, these attributes are not kept constant, meaning that the implicit comparison is now one between the job satisfaction associated with a "typical" job in the public sector and the job satisfaction associated with a "typical" job in the private sector, that is, *mutatis mutandis*.

4.1. German Socio-Economic Panel

The data have been extracted from the German Socio-Economic Panel, 2004. We base our analysis on that particular year because it includes a relatively rich menu of questions that are potentially related to a person's preferences for public and private sector employment. These questions were not included in other years of the survey. Our sample and variable selection follows in part the prior study of Dustmann and Van Soest (1998) who studied self-selection in a model for public and private sector wages. We focus on male workers and use the same instruments for sector choice as they did, namely the father's occupational status (white collar, civil servant) when the worker was 15, as well as the mother's employment status at that age.

In contrast to Dustmann and Van Soest, we do not include the entire working age population but focus on younger workers, those aged between 25 and 40. The reason is that, when modeling the effect of preference heterogeneity on choice, one ideally would like to observe these preferences at the time of choice. Over time, they can change and the interpretation of measured correlations as being related to self-selection based on preference heterogeneity becomes more and more difficult, in particular, as many workers are locked in their sector and cannot adjust to preference changes because switching costs are high. While it might be the case that preferences systematically adapt in order to rationalize a choice ex post (e.g., to avoid cognitive dissonance), thus strengthening measured correlations, they might as well evolve in ways altogether unrelated to the choice. Unfortunately, we cannot observe choice-moment preference variables in our data. However, we can reduce the problem by considering young workers relatively soon after their sector choice at the beginning of their careers.

Table 1 presents variable definitions and means (with their standard errors in parentheses) for the sample of 1,756 observations, separately by sector. Average job satisfaction is slightly higher in the public sector (7.2 relative to 7.1), but the difference is not statistically significant. Private sector earnings are about 8% higher on average, a statistically significant difference.

Table 1. Variable Definitions and Means by Sector.

Variable	Definition	Mean (SE)	
		Public	Private
JOB SATISFACTION	Coded on a 0, 1, …, 10 scale	7.208	7.135
		(0.107)	(0.051)
GERMAN	Citizenship (yes = 1)	0.952	0.865
		(0.012)	(0.009)
MARRIED	Marital status (yes = 1)	0.502	0.584
		(0.028)	(0.013)
MEDIUM FIRM	Firm has more than 100 workers	0.356	0.294
		(0.026)	(0.012)
LARGE FIRM	Firm has more than 2,000 workers	0.450	0.225
		(0.027)	(0.011)
EDUCATION	Years of formal schooling	13.4	12.4
		(0.155)	(0.071)
WORKING HOURS	Weekly regular hours	42.7	44.1
		(0.489)	(0.253)
OVERTIME	Weekly overtime hours	2.889	2.7
		(0.248)	(0.106)
LOG EARNINGS	Logarithm of current monthly gross labor income (in Euro)	7.809	7.884
		(0.030)	(0.015)
AGE	Age (in years)	34.2	34.2
		(0.242)	(0.114)
POOR HEALTH	A caseness score between 0 (perfect health) and 8 (poor health)	1.269	1.242
		(0.106)	(0.051)
HELP	Importance of being there for others (very important/important = 1)	0.894	0.914
		(0.017)	(0.007)
SUCCESS	Importance of being successful in ones career (very important/important = 1)	0.792	0.806
		(0.022)	(0.010)
ENGAGEMENT	Importance of political and social engagement (very important/ important = 1)	0.353	0.234
		(0.026)	(0.011)
RISK	Willingness to take risks (0 = "none"; 10 = "full")	5.314	5.333
		(0.117)	(0.056)
F. WHITE COLLAR	Occupational status of father at age 15	0.251	0.215
		(0.024)	(0.011)
F. CIVIL SERVANT	Occupational status of father at age 15	0.178	0.072
		(0.021)	(0.007)
M. EMPLOYED	Employment status of mother at age 15	0.239	0.242
		(0.023)	(0.011)
OBSERVATIONS		331	1,425

Among the standard socio-economic controls, AGE, EDUCATION, MARRIED, and POOR HEALTH, only the last deserves additional comment as it is an ''objective'' measure of poor health, a caseness score. It is based on the following eight indicators: Frequency (always/often/sometimes = 1) of strong physical pains; underachievement or limitations at work or during everyday tasks due to physical health problems; underachievement or limitations due to physical health problems; social limitations due to impaired health; affect of state of health (greatly/slightly = 1) on climbing stairs; affect of state of health on other tiring everyday tasks.

In addition, we observe a number of preference indicators regarding risk, social responsibility, and career orientation. In 2004, survey participants were asked about the importance they place on the following three aspects of life: having a successful career (SUCCESS); helping other people (HELP); being engaged in social and political activities (ENGAGEMENT). The important questions were asked on a four-point scale, with responses ''unimportant/not very important/important/very important,'' and we define dummy variables taking the value 1 for outcome ''important'' or ''very important.'' The risk variable is also a self-assessment, measured on an 0–10 scale (''How do you see yourself: are you a person who is fully prepared to take risks, or do you try to avoid taking risks?''). Our conjecture was that career-oriented individuals and those willing to take higher risks are more likely to be found in the private sector, whereas individuals who put more importance on helping and public service tend to be matched to the public sector. From Table 1, however, only the incidence of ENGAGEMENT differs statistically significantly between the two sectors.

4.2. Results

A total of six models were estimated, two each using the independence copula, the normal copula, and the Frank copula, respectively. In Model 1, the regressors in the outcome equation include GERMAN, MARRIED, EDUCATION, AGE, POOR HEALTH, HELP, SUCCESS, ENGAGEMENT, and RISK. The selection equation includes the same variables plus three instruments, FATHER WHITE COLLAR, FATHER CIVIL SERVANT, MOTHER EMPLOYED, all dummy variables. In Model 2, five job-specific attributes were added, namely MEDIUM FIRM, LARGE FIRM, WORKING HOURS, OVERTIME, LOG EARNINGS.

Table 2 shows the log-likelihood values and the correlation parameters for these models. There is clear evidence against the null hypothesis of

Table 2. Log-Likelihood and Estimated Dependence Parameters.

Copula	Model 1	Model 2
Independence		
Log likelihood	−4084.8	−4069.3
Normal		
ρ_1	0.3191	0.3094
	(0.422)	(0.435)
ρ_0	−0.6842	−0.7133
	(0.129)	(0.114)
Log likelihood	−4081.0	−4064.5
Frank		
θ_1	−1.1381	−0.9485
	(2.119)	(2.127)
θ_0	−5.0381	−5.7781
	(1.693)	(1.691)
Log likelihood	−4080.3	−4063.3

Note: Standard errors in parentheses; job-specific attributes are excluded in Model 1 but included in Model 2.

random selection of workers into the two sectors. There are four possible comparisons, independence against normal copula and independence against Frank copula, for Model 1 and Model 2. A likelihood ratio test rejects the independence model in all four cases. The test statistic varies between 7.6 and 12.0, with critical 5% value for 2 restrictions of 5.99. A likelihood comparison of the normal copula and the Frank copula favors the latter, although the difference is just 0.7 in Model 1 and 1.2 in Model 2. The horizontal comparison between Model 1 and Model 2 shows that the job attributes are jointly significant indeed. However, as pointed out earlier, the comparison between Model 1 and Model 2 should be made based on the type of interpretation one wants to attach to the public/private sector comparison rather than on statistical grounds.

Substantively, the two models agree with regards to self-selection patterns. The nature of the selection process can be inferred from the estimates of ρ_1, ρ_0, θ_1, and θ_0. Recall that ρ_1 and θ_1 model dependence between sector choice and public sector job satisfaction, whereas ρ_0 and θ_0 model dependence between sector choice and private sector job satisfaction. In both Frank and normal copula, negative values indicate that the two random variables, ε_s and v, for $s = 0, 1$, tend to move in opposite direction. A value of zero represents independence, while positive values arise from comovements.

From Table 2, one cannot reject that selection into the public sector is independent of public sector job satisfaction, meaning that the job satisfaction distribution of those who work in the public sector does not differ from the distribution of an arbitrary worker with the same observed characteristics. In contrast, the private sector selection parameters ρ_0 and θ_0 are negative and significant. The Spearman rank correlations implied by the estimates for θ_0 are -0.62 in Model 1, and -0.67 in Model 2, respectively. The negative correlations mean that the private sector counterfactual job satisfaction of those who actually opted for the public sector is below than that of an average worker. Taken together, these two observations provide some evidence of ''optimal'' self-selection based on unobservables: By working in the public sector, public sector types are better off, since they avoid the below average job satisfaction they would receive from a private sector job.

Table 3 contains the regression coefficients for the normal and Frank copula estimates of Model 2. The first three columns show the estimated regression parameters for the normal copula (public sector job satisfaction, private sector equation, and selection equation). The estimated parameters for the Frank copula follow in the next three columns. The threshold parameters are available on request.

The most conspicuous aspect of Table 3 is the stability of the estimates across specification, corroborating the similarity of the normal and Frank results found in Table 2. Differences between the normal and the Frank regression parameters are small and often restricted to the second or third decimal place. The additional gain from having introduced the copula framework, for this particular application, is thus primarily the insight that the results are robust to modeling dependence by either a normal or Frank copula, which was not to be expected ex ante.

As to the substantive results, we find significant positive effects of being German, being not married and having a higher education on the probability of working in the public sector. Moreover, those who find it important or very important to show civic engagement are more likely to work in the public sector. As typically found in the literature, the job satisfaction index is *u*-shaped in age (*ceteris paribus*, controlling for health and other factors that also vary with age) and poor health reduces job satisfaction.

Sector-specific differences are found for earnings, education, overtime work, and marital status. The point estimates for the effect of earnings on job satisfaction is positive in both sectors, but the effect is almost twice as large, and statistically significant only, in the private sector. Job satisfaction

Table 3. Self-Selection Ordered Probit Models of Sector-Specific Job Satisfaction (German Socio-Economic Panel 2004, $N = 1{,}756$).

	Normal Copula			Frank Copula		
	Public	Private	Selection	Public	Private	Selection
MEDIUM FIRM	0.0595	0.0390		0.0604	0.0335	
	(0.160)	(0.060)		(0.156)	(0.058)	
LARGE FIRM	0.0329	−0.0037		0.0261	0.0001	
	(0.160)	(0.069)		(0.160)	(0.068)	
WORKING HOURS	0.0071	0.0001		0.0076	0.0000	
	(0.009)	(0.003)		(0.009)	(0.003)	
OVERTIME	−0.0285*	0.0018		−0.0283*	0.0033	
	(0.017)	(0.007)		(0.017)	(0.007)	
LOG EARNINGS	0.1467	0.2531*		0.1275	0.2406*	
	(0.140)	(0.055)		(0.139)	(0.055)	
GERMAN	−0.0882	0.1253	0.4174*	0.0788	0.1150	0.4047*
	(0.328)	(0.085)	(0.139)	(0.371)	(0.086)	(0.140)
MARRIED	−0.0493	0.1181*	−0.1497*	−0.0717	0.1136*	−0.1484*
	(0.143)	(0.062)	(0.079)	(0.148)	(0.060)	(0.080)
EDUCATION	0.0036	−0.0313*	0.0554*	0.0241	−0.0293*	0.0550*
	(0.026)	(0.011)	(0.014)	(0.032)	(0.011)	(0.014)
AGE	−0.3510	−0.2490*	−0.0949	−0.3857*	−0.2614*	−0.1004
	(0.225)	(0.105)	(0.135)	(2.204)	(1.036)	(1.364)
AGE SQUARED	0.5008	0.3640*	0.1452	0.5562*	0.3837*	0.1519
	(0.338)	(0.158)	(0.203)	(0.331)	(0.155)	(0.204)
POOR HEALTH	−0.1739*	−0.1675*	0.0158	−0.1667*	−0.1623*	0.0199
	(0.034)	(0.016)	(0.020)	(0.039)	(0.017)	(0.020)
HELP	0.3139	0.2104*	−0.1009	0.2950	0.2163*	−0.1104
	(0.214)	(0.094)	(0.118)	(0.224)	(0.091)	(0.120)
SUCCESS	0.1550	0.0989	−0.1155	0.0927	0.0792	−0.1226
	(0.146)	(0.073)	(0.098)	(0.153)	(0.073)	(0.100)
ENGAGEMENT	−0.0192	−0.0055	0.2587*	0.0643	−0.0138	0.2759*
	(0.140)	(0.067)	(0.081)	(0.154)	(0.065)	(0.081)
RISK	−0.0047	0.0200	−0.0124	−0.0074	0.0203*	−0.0119
	(0.033)	(0.012)	(0.018)	(0.031)	(0.012)	(0.018)
F. WHITE COLLAR			0.0865			0.0721
			(0.083)			(0.085)
F. CIVIL SERVANT			0.4935*			0.4900*
			(0.109)			(0.110)
M. EMPLOYED			−0.1246			−0.1365*
			(0.078)			(0.079)

*indicates statistical significance at the 10% level.

falls with years of formal education in the private sector, while working overtime hours has a significant negative effect on job satisfaction only in the public sector.

To obtain a sense for the magnitude of these effects, one could convert the implied index changes into changes in predicted probabilities. An alternative, and much simpler, possibility for interpreting the coefficients is to look at relative magnitudes, that is, at trade-off ratios. For example, the estimated coefficient of being married in the private sector is of opposite sign and about two thirds of the absolute value of the health coefficient. Thus, being married rather than single compensates (in the sense of keeping the job satisfaction distribution unchanged) for a two-third point (or one-third standard deviation) increase in the health caseness score, reflecting the substantial importance of health for job satisfaction.

5. CONCLUSIONS

The methodological developments in the chapter were motivated by a substantive issue related to job satisfaction. Job satisfaction is an important economic outcome. More satisfied workers are less likely to quit. Among older workers, those who are more content with work are less likely to retire.

In this chapter, we have proposed to study the determinants of job satisfaction using a new class of ordered probit models with self-selection. The class has two main features: First, it preserves marginal probit distributions for the ordered outcome and binary selection models, and thus generalizes the standard econometric model without self-selection. Second, it accounts for the joint determination of outcome and selection in a simple, yet flexible parametric framework. Thus, implementation of these methods does not require any estimation and inferential methods beyond those of maximum likelihood. In this sense, our chapter offers an alternative to other recent implementations of switching regression models for ordered responses based on joint normality (DeVaro, 2006; Munkin & Trivedi, 2008).

Using a sample of young German men from the German Socio-Economic Panel, we could reject the null hypothesis of independence between job satisfaction and sector choice. In particular, we found evidence of "optimal" self-selection based on unobservables: By working in the public sector, public sector types are better off, since they avoid the below average job satisfaction they would receive from a private sector job. It turned out that the conclusions were robust to the choice of copula, as long as dependence was allowed for. From a computational point of view, the model based on

the Frank copula avoids numerical integration and is easier to maximize. In our applications, computation time was cut by about two thirds.

Ordered response models with endogenous switching, as discussed in this chapter, have applications in many other areas of empirical economics. Future research should pursue some obvious extensions of these methods, including an integration of additional copula functions beyond the three considered in this chapter, and more general, multinomial selection mechanisms. In subjective well-being research, the endogeneity of choice variables should be addressed more carefully. The methods proposed in this chapter provide a framework for doing so.

ACKNOWLEDGMENT

We thank Murray Smith as well as three anonymous referees for valuable comments on an earlier version of the paper.

REFERENCES

Asiedu, K. F., & Folmer, H. (2007). Does privatization improve job satisfaction? The case of Ghana. *World Development*, *35*, 1779–1795.

Boes, S., & Winkelmann, R. (2006). Ordered response models. *Advances in Statistical Analysis*, *90*(1), 165–180.

Clark, A. (1997). Job satisfaction and gender: Why are women so happy at work? *Labour Economics*, *4*, 341–372.

DeVaro, J. (2006). Teams, autonomy, and the financial performance of firms. *Industrial Relations*, *45*, 217–269.

Dustmann, C., & Van Soest, A. (1998). Public and private sector wages of male workers in Germany. *European Economic Review*, *42*, 1417–1441.

Greene, W. H., & Hensher, D. A. (2008). *Modeling ordered choices: A primer and recent developments*. Working Paper no. 08-26. Department of Economics, Stern School of Business.

Heywood, J. S., Siebert, W. S., & Wei, X. (2002). Worker sorting and job satisfaction: The case of union and government jobs. *Industrial and Labor Relations Review*, *55*, 596–610.

Lee, L. (1983). Generalized econometric models with selectivity. *Econometrica*, *51*, 507–512.

McCausland, W. D., Pouliakas, K., & Theodossiou, I. (2005). Some are punished and some are rewarded: A study of the impact of performance pay on job satisfaction. *International Journal of Manpower*, *26*, 636–659.

Munkin, M. K., & Trivedi, P. K. (2008). Bayesian analysis of the ordered probit model with endogenous selection. *Journal of Econometrics*, *143*, 334–348.

Nelson, R. B. (2006). *An introduction to copulas*. Berlin: Springer.

Pederson, P. J., Schmidt-Sorensen, J. B., Smith, N., & Westergard-Nielsen, N. (1990). Wage differentials between the public and private sectors. *Journal of Public Economics*, *41*, 125–145.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, *8*, 229–231.

Smith, M. D. (2003). Modeling sample selection using Archimedean copulas. *Econometrics Journal*, *6*, 99–123.

Smith, M. D. (2005). Using copulas to model switching regimes with an application to child labour. *The Economic Record*, *81*, S47–S57.

Trivedi, P. K., & Zimmer, D. M. (2007). Copula modeling: An introduction for practitioners. *Foundations and Trends in Econometrics*, *1*, 1–111.

van der Gaag, J., & Vijverberg, W. P. M. (1988). A switching regression model for wage determinants in the public and private sectors of a developing country. *Review of Economics and Statistics*, *70*, 244–252.

van Ophem, H. (1999). A general method to estimate correlated discrete random variables. *Econometric Theory*, *15*, 228–237.

van Ophem, H. (2000). Modeling selectivity in count data models. *Journal of Business and Economic Statistics*, *18*, 503–511.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypothesis. *Econometrica*, *57*(2), 307–333.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, *50*, 125.

Zimmer, D. M., & Trivedi, P. K. (2006). Using trivariate copulas to model sample selection and treatment effects: Application to family health care demand. *Journal of Business and Economic Statistics*, *24*, 63–76.

Zweimüller, J., & Winter-Ebmer, R. (1994). Gender wage differentials in private and public sector jobs. *Journal of Population Economics*, *7*, 271–285.

THE SURVIVAL AND GROWTH OF ESTABLISHMENTS: DOES GENDER SEGREGATION MATTER?

Helena Persson and Gabriella Sjögren Lindquist

ABSTRACT

We empirically study gender segregation in privately owned Swedish establishments, and the correlation between gender segregation, survival and growth of establishments. We find that the overall inter-establishment gender segregation in Sweden has been constant between 1987 and 1995 and at the same level as that found in US manufacturing. Our results show that establishments dominated by males or females have a higher probability of exiting the market than more integrated establishments and that establishments dominated by females grow more slowly than other establishments. An important additional finding is that establishments with a skewed workforce in terms of educational background have lower survival probabilities. Furthermore, establishments with skewed age distributions have both lower survival probabilities and grow less compared with other establishments. These findings are consistent with theories suggesting that workers with different demographic characteristics contribute to a creative working environment as a result of their different experiences, a greater variety of information sources and different 'thinking'.

Jobs, Training and Worker Well-Being
Research in Labor Economics, Volume 30, 253–282
Copyright © 2010 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 0147-9121/doi:10.1108/S0147-9121(2010)0000030011

1. INTRODUCTION

Gender segregation prevails in all labour markets and is regarded as a problem by the population at large as well as by representative and legislative bodies alike. Evidence of this concern can be seen in the ongoing debate over gender equality in the labour market and in the political ambition to reduce the segregation in both the internal and external labour markets. The labour markets in many countries are regulated by equal opportunity laws and affirmative action plans, that attempt to force them to become more integrated.[1]

Research in the field of economics addressing inter-firm gender segregation has focused mainly on the extent of gender segregation between firms and the relation between the gender distribution of the workers and the gender wage gap. The point of departure of these earlier studies is the model of employer-taste discrimination in Becker (1957).[2] An extension to these previous studies is the study by Hellerstein, Neumark, and Troske (2002), who test the long-run implications of the Becker model on firm profits and firm growth. According to Becker's model, firms that employ a large fraction of women will be relatively more profitable due to lower wage costs, and thus enjoy a greater probability of growing by underselling other firms in the competitive product market.[3] Hellerstein et al. look at the way gender segregation affects firms' profits in the US and examine whether firms employing a large share of women actually expand more, implicitly as an effect of lower wage costs.[4] They find clear evidence of a positive relationship between profit and the proportion of female workers among firms with market power but no evidence of that firms that employ a large share of women expand.

In contrast to previous studies that focused on the gender wage gap, we focus on the dynamics of gender segregation and its correlation to establishment survival and growth. Different theoretical models give different implications about gender segregation, employment dynamics and firm profitability. Like Becker's 'taste'-based model of discrimination, Lang's (1986) language model implies firm segregation in the long run. Lang develops a model in which people can only work together if they 'speak' the same language and in which it is costly to learn a second language. Language refers to all aspects of verbal and non-verbal communication. Blacks and whites or men and women can be said to 'speak' different languages in this sense. The competitive market will tend to minimize communication costs through segregation.

Mello and Ruckes (2006) present a model of team composition where heterogeneous teams have greater variety of information sources than

homogenous teams and thereby reach better decisions, given that information and preferences can be expressed openly. However, members of heterogeneous teams are more likely to diverge in their preferences with respect to courses of action, which is reflected in lower effort. The model predicts that it will be more profitable for the firm to have a heterogeneous workforce in dynamic and uncertain situations, while homogenous teams are preferred when there is little decision uncertainty.

In the corporate governance literature, gender diversity and firm performance has been studied to a quite large extent and both the theoretical and empirical evidence are ambiguous. Among arguments working for diversity are that a more heterogeneous board bases its decisions on more alternatives compared to a more homogenous board. The board quality may also become higher when board members can be chosen from a larger group of candidates when women are also included as potential members. Arguments against management diversity are that heterogeneous boards probably experience more conflicts and are more time-consuming due to more opinions (see Smith, Smith, & Verner, 2006).

In this study, we examine gender segregation at the establishment level, and the correlation between establishments' survival and growth. We first analyse the gender distribution of employees at the establishment level and examine the way it changes over time. We then examine whether there is a correlation, and whether it is positive or negative, between the gender distribution of the establishments and establishment survival and growth. For this aim, we use a unique matched employer–employee dataset consisting of all Swedish privately owned establishments. Further, we examine mature and new establishments separately since there are conspicuous differences between mature and new establishments in terms of their probability of survival and/or potential for growth. We also believe that new establishments can affect their own gender distribution to a greater extent than mature establishments, since the former chose their workforce at the start up.

The results indicate that overall segregation has not changed between 1987 and 1995. The extent of gender segregation between establishments in Sweden is comparable to that found in US manufacturing, but less than in Portugal or Korea (the only two countries besides the US providing comparable information). However, we find that Swedish establishments with a moderate male bias (i.e., firms with 50–75 per cent male employees) become more segregated over time, while all other establishments are becoming more integrated and that new establishments are as segregated as mature ones.

We also find that gender-segregated establishments have a higher risk to exit the market. That is, both female- and male-dominated establishments have lower survival rates than establishments with a more even gender distribution. However, female-dominated establishments have a lower growth rate than more integrated and male-dominated establishments. Additionally, we find that establishments that are heterogeneous with respect to gender, age and education seem to be more successful in terms of survival and growth than more homogeneous establishments. Hence, our empirical results are in line with theories suggesting that heterogeneous work compositions promote higher firm payoffs.

The chapter is organized as follows. The data are presented in Section 2. In Section 3 gender segregation is measured. In Section 4 establishment survival and gender segregation are analysed, and in Section 5 we study establishment growth and gender segregation. Section 6 offers some concluding comments.

2. THE LINKED ESTABLISHMENT–EMPLOYEE DATA

In order to examine the gender distribution and its relation to the survival and growth of establishments we use a dataset consisting of all privately owned establishments[5] in Sweden between 1986 and 1995 and all workers employed in these establishments.[6]

The following is a brief description of the method of producing the data.[7] Information about all workers aged 16–64 have been taken from the Swedish Employment Register, which covers the whole population aged 16 or more, in November each year. The connection between the employer and the employee or self-employed is denoted by the identity numbers of the firms and the establishments where each individual had his or her main work. These identity numbers are taken from the Business Register, in which every firm and every establishment is assigned a unique such number. However, the identity number may change due to an alteration in the relevant legal form or due to error. This is especially likely to occur in the case of small establishments. Since most of the new or closed-down establishments are small, it is not enough simply to measure changes in the identity number from the Business Register in order to define these categories. By also noting individuals associated with establishments over consecutive years, it becomes possible to distinguish 'true' births and 'true' deaths from what are in fact only changes in a unit's identity.

The identity variable in the dataset is the establishment. In the case of establishment deaths before 1995, information is provided on the year of death. For each year we also use information about the number of employees, the relevant industry and whether the establishment is part of a multi-unit firm. The employees at each establishment are also further disaggregated into gender groups and four age groups, as well as four educational groups based on the level of education attained. The age and educational groups are further divided into their own gender groups. The data allow us to follow the worker distribution over several years.[8]

As mentioned above, we examine mature and new establishments separately. Hence, we use two different samples in our analysis. The *first sample* consists of privately owned establishments that existed in 1986, which are defined as 'mature' establishments. A drawback of the data for mature establishments is that it is left truncated. The first year possible to create employer–employee data from Swedish statistics is 1985. This means that there is no information on how long an establishment has existed if it was established before 1985. Since data of 1985 contain a lot of teething problems, we choose 1986 as our starting year. Hence, the results for the mature establishments might be biased since it is based in the stock of existing establishments at a certain year with different elapsed duration. To the extent that establishments with different worker characteristics also have systematically different distributions of elapsed duration, by not conditioning on the elapsed duration might bias the results.

Since we are interested in changes in gender distribution, the establishment should not be too small. Therefore, we include only establishments with at least four employees in 1988. The establishments are then followed up over the period 1988–1995. In 1988, 64,005 mature establishments employed at least four workers.

The *second sample* consists of all privately owned new establishments that were created in Sweden in 1987 and 1988 and had at least four employees in the second year of their existence.[9] They are followed up until 1994 and 1995, respectively. Most of the new establishments are very small. Around 50 per cent of them employed only one person in their second year. The second sample thus consists of 9,543 new establishments.[10]

Table 1 shows the number of establishments existing in each year, the proportion of them that survives over the period studied, and the number of their employees. After eight years, 76 per cent of the mature establishments and 55 per cent of the new establishments were still in existence. The mature establishments are larger than the new ones, employing on average 22–26 individuals compared to 11–14 among the new ones.

Table 1. Survival and Employment in Privately Owned Establishments with at least Four Employees.

	1988	1989	1990	1991	1992	1993	1994	1995
Number of survivors								
Mature establishments	64,005	62,530	60,109	57,309	54,192	51,096	48,627	–
Cohort 1987	4,667	4,318	3,910	3,588	3,196	2,834	2,614	–
Cohort 1988	–	4,876	4,478	4,082	3,554	3,125	2,842	2,633
Per cent survivors								
Mature establishments	100	98	94	90	85	80	76	–
Cohort 1987	100	93	84	77	68	61	56	–
Cohort 1988	–	100	92	84	73	64	58	54
Average number of employees								
Mature establishments	25.5	25.2	24.7	24.3	23.0	21.9	23.2	–
Cohort 1987	11.7	11.8	12.2	12.1	11.9	11.7	12.4	–
Cohort 1988	–	11.9	11.9	11.6	11.5	11.4	12.6	13.9

Notes: All mature establishments had at least four employees in 1988. All new establishments had at least four employees in the second year of their existence (in 1988 and 1989, respectively).

Table 2. Descriptive Statistics, First Year Observations.

	Mature	New
Per cent multi-units	31	19
Per cent men	62	59
Per cent aged 16–24	20	25
Per cent aged 25–54	67	67
Per cent aged 55–59	7	5
Per cent aged 60–64	5	3
Per cent pre-upper secondary school	38	30
Per cent secondary school	45	48
Per cent university <3 years	6	8
Per cent university ≥ 3 years	5	7
Missing information on education	6	7
Number of establishments	64,005	9,543

Notes: All mature establishments had at least four employees in 1988. All new establishments had at least four employees in the second year of their existence (in 1988 and 1989, respectively).

Table 2 provides descriptive statistics for all mature and new privately owned establishments. The statistics for mature establishments refer to 1988, while those for new establishments refer to 1988 in the case of 1987 start-ups and to 1989 in the case of 1988 start-ups. The distribution of

workers between new and mature establishments is quite similar. On average, workers are a bit younger and more educated in new establishments. The main difference between new and mature establishments is that a larger part of the mature establishments than the new, 31 compared to 19 per cent, are 'multi-units', i.e., are part of a larger firm with more than one establishment.

3. MEASURING GENDER SEGREGATION

We start by analysing the gender distribution of employees at the establishment level. Average segregation within establishments and the way it develops over time can be described with the help of segregation indexes. The most commonly used measures of segregation are the Duncan and Duncan dissimilarity index and the Gini coefficient (see, e.g., Carrington & Troske, 1997). These indexes measure the extent to which the distribution on men and women across establishments deviates from an even distribution whereby each group is proportionally represented in each establishment.

3.1. The Duncan and Duncan Dissimilarity Index and the Gini Coefficient

We estimate gender segregation between establishments in Sweden by (i) the Duncan and Duncan dissimilarity index, defined as

$$D = \sum_{i=1}^{T} \frac{1}{2} |w_i - m_i| \tag{1}$$

where w_i and m_i are establishment i's share of female and male employees in the sample used, $D \in [0, 1]$ where 0 equals total integration and 1 equals total segregation, and (ii) the Gini coefficient of segregation,[11] defined as

$$G = 1 - \sum_{i=1}^{T} w_i \left(m_i + 2 \sum_{j=i+1}^{T} m_j \right) \tag{2}$$

where T is the number of establishments in the sample.

The establishments are sorted on the basis of w_i/m_i where the establishment with the lowest share of female workers is ranked number 1 and the establishment with the largest fraction is ranked T. $G \in [0, 1]$, where 0 equals total integration and 1 equals total segregation.

An important problem about these indexes is that if we have only few employees in each establishment, the indexes may indicate the existence of segregation even if workers are allocated randomly across units. Consider a large sample of two-person firms, which, taken together, employ a 50/50 mix of men and women. Random allocation of workers to firms will result in 25 per cent of the firms employing two men, 50 per cent of employing one man and one woman and 25 per cent employing two women. Both segregation indexes would report substantial segregation in this case. The Duncan and Duncan index and the Gini coefficient would equal 0.50 and 0.75, respectively. From this we can see that the Gini coefficient is even more sensitive to random allocation than the dissimilarity index. If instead we consider a sample of firms employing 1,000 workers and randomly allocate a 50/50 mix of men and women, we end up with a Duncan and Duncan index and a Gini coefficient of 0.03 and 0.01, respectively. Hence, small firms tend to inflate the indexes.

Carrington and Troske (1997) have proposed modifications of the Gini coefficient as a means of distinguishing between systematic and random segregation. The Gini coefficient of systematic segregation is defined as

$$\hat{G} = \frac{G - G^*}{1 - G^*} \text{ if } G - G^* \geq 0 \text{ and } \hat{G} = \frac{G - G^*}{G^*} \text{ if } G - G^* < 0 \qquad (3)$$

where $\hat{G} \in [-1, 1]$.

G is the standard Gini coefficient stated above and G^* is the Gini coefficient that would occur if a very large number of workers were allocated randomly to employers, taking each gender's share of the population and the size distribution of the establishments as determined by the sample. If the Gini coefficient of systematic segregation equals 0, then the gender distribution is totally random. If the coefficient is larger than 0, there is systematic gender segregation.

The Duncan and Duncan dissimilarity index is analogously modified. D is the standard dissimilarity index and D^* is the average dissimilarity index obtained if workers are assigned randomly to establishments.

The segregation indexes for all Swedish privately owned establishments are reported in Table 3. Both the Duncan and Duncan dissimilarity index and the Gini coefficient (see columns 1, 3, 5 and 7 in Table 3) indicate that new establishments are more segregated than mature establishments.

However, new establishments are smaller than mature establishments, which, as noted above, tend to inflate the indexes. Controlling for the random allocation of workers by using systematic indexes (see columns 2, 4, 6 and 8 in Table 3), we find no difference in gender segregation between new

Table 3. Gender Segregation.

Year	Mature Firms				New Firms			
	Dissimilarity index [1]	Systematic dissimilarity index [2]	Gini coefficient [3]	Systematic Gini coefficient [4]	Dissimilarity index [5]	Systematic dissimilarity index [6]	Gini coefficient [7]	Systematic Gini coefficient [8]
2	0.447	0.367	0.602	0.502	0.501	0.370	0.671	0.520
3	0.446	0.366	0.599	0.499	0.496	0.366	0.665	0.514
4	0.445	0.364	0.597	0.496	0.494	0.365	0.661	0.509
5	0.442	0.359	0.594	0.490	0.493	0.363	0.661	0.508
6	0.443	0.357	0.596	0.489	0.487	0.355	0.654	0.496
7	0.448	0.361	0.602	0.493	0.492	0.366	0.657	0.507
8	0.448	0.364	0.601	0.496	0.491	0.371	0.656	0.514
No. of establishments in year 2	64,005	64,005	64,005	64,005	9,543	9,543	9,543	9,543
No. of establishments in year 8	48,627	48,627	48,627	48,627	5,247	5,247	5,247	5,247

and mature firms.[12] That is, new establishments are *not* more segregated than mature ones.

The indexes are also relatively stable over time for all establishments. This could mean either that establishments maintain their gender distribution over the years, or that the more segregated (integrated) establishments exit the market while those that survive become more segregated (integrated). To explore this, we therefore made separate analyses of establishments that survived all eight years (not shown here). However, the results were very similar to those given in Table 3 above, which indicates that establishments do maintain their gender distribution over the years. This will be studied further in Section 3.3.

3.2. Gender Segregation in Other Countries

Table 4 illustrates gender segregation indexes among establishments in the US, Portugal and Korea. Establishments in Korea are highly gender segregated but the segregation has been declining over the past 30 years. The Portuguese establishments are also highly segregated, but the gender distribution between establishments has remained stable over the past 14 years. The Swedish gender segregation is on the same level as in US manufacturing, but below that in Korea and Portugal. Also, as in Portugal, it is stable over time.

3.3. Changes in Gender Distribution within Establishments

Previous sections do not provide information on the change in gender distribution within the individual establishments. For example, segregation can increase in the labour market either because establishments that are already segregated become more so, or because integrated establishments change their gender distribution and become more segregated. Alternatively, some establishments may change the composition of their workforce from domination by one gender to domination by the other. These possibilities will be examined further in this section.

In Fig. 1, in order to examine changes in gender distribution in the individual establishments, we have plotted the gender distribution in mature and new establishments in the first and last years observed. Each plot represents an establishment. The x-axis shows the establishments' share of men the first year observed, while the y-axis shows the same establishments'

Table 4. Segregation Indexes for Different Countries.

Study	Description	Dissimilarity Index	Systematic Dissimilarity Index	Gini Coefficient	Systematic Gini Coefficient
Cabral Viera et al. (2005)	Portuguese privately owned establishments in manufacturing, agriculture and service sectors. 1985–1999[a]				
	1985	0.553	0.492	0.732	0.670
	1999	0.563	0.489	0.742	0.668
Carrington and Troske (1998)	US manufacturing, 1990	0.43	0.33	0.59	0.45
Carrington and Troske (1995)	US establishments with <35 shareholders and <100 employees, 1982	0.66[b]			
Yoon et al. (2003)	Korean establishments with ≥ 10 employees in all sectors except agriculture, forestry, fishing, public administration, educational and medical services. 1971–1998[c]				
	1971	0.65	0.49	0.83	0.69
	1986	0.59	0.48	0.75	0.67
	1998	0.52	0.39	0.70	0.54

[a]Information on all years between 1985 and 1999 can be found in Cabral Viera et al. (2005).
[b]The authors do not calculate any systematic segregation indexes but compare their results with the gender distribution implied by a random hiring model. Using a χ^2-test they reject the random hiring model at the 95 per cent level and conclude that there is gender segregation not stemming from random allocation.
[c]Information on almost all years between 1971 and 1998 can be found in Yoon et al. (2003).

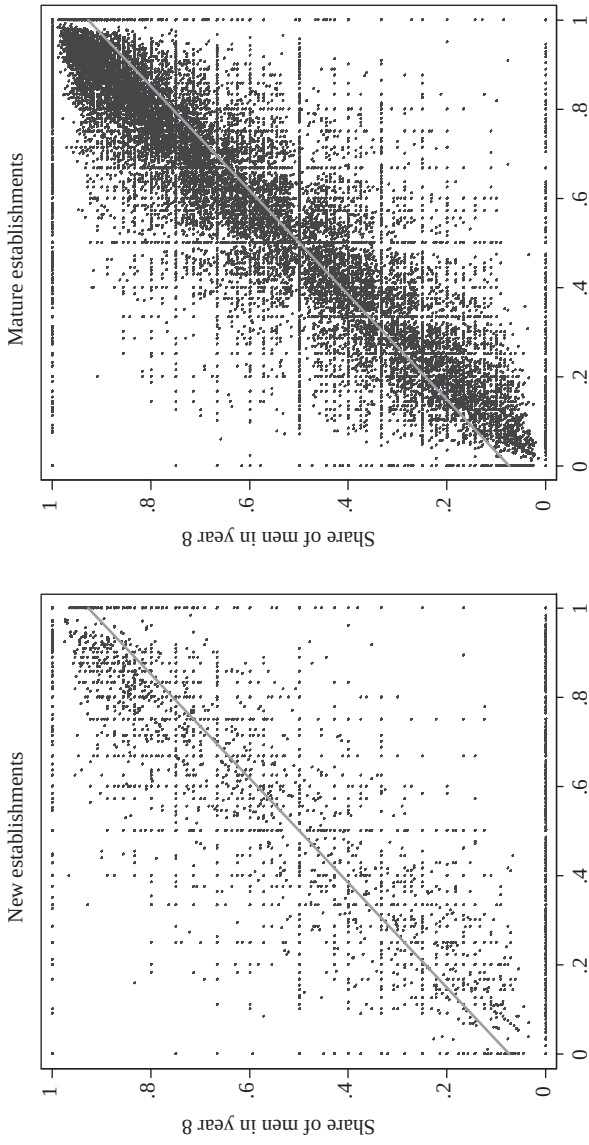

Fig. 1. The Change in the Share of Male Workers in Establishments between Years 2 and 8.

share of men the last year observed. A dot along the diagonal means that the establishment has the same share of men the first and the last year observed. These figures reveal a considerable variety in the gender distribution, and in the way this evolved in the individual establishments, over the seven-year period examined: in some establishments it did not change, while in others it did. However, the plotting shows most establishments in a wide band along the diagonal, which suggests that the various establishments chose widely differing gender compositions but that within each establishment the gender distribution has changed very little seven years later. This conclusion is particularly valid for mature establishments. The same observation occurs in Haltiwanger, Lane, and Spletzer (2007), which among other things study the choice of gender composition in new firms. The authors conclude that firms choose their workforce quite deliberately.

Table 5 shows the (mature and new) establishments' percentage point change in their share of male workers between the first and last years observed. These results confirm the interpretation of Fig. 1 that the gender distribution within establishments changes little over time. At both these points in time, 17 per cent of the mature and 21 per cent of the new establishments have the same gender distributions.

We also separate establishments into three size groups: establishments employing less than 10 workers, 10 or more workers and 20 workers or more. The last group is a sub-sample of the second group. According to size, we see that as many as 28 per cent of the new and 25 per cent of the mature establishments employing less than 10 employees maintain their gender distribution over time. Only 2 per cent of both new and mature

Table 5. Percentage Point Change in Establishments' Gender Distribution between the First and Last Years of Observation.

Percentage Point Change, x	All Mature Establishments	Mature Establishments Employing			All New Establishments	New Establishments Employing		
		<10	≥ 10	≥ 20		<10	≥ 10	≥ 20
$x = 0$	17	25	6	2	21	28	7	2
$0 < x \leq 10$	43	26	68	79	33	20	56	69
$10 < x \leq 25$	26	28	23	17	26	26	28	22
$25 < x \leq 50$	11	17	3	2	16	19	8	6
$50 < x \leq 75$	2	3	0	0	3	5	1	1
$75 < x$	1	1	0	0	1	2	0	0
No. of establishments	48,627	27,905	20,722	10,523	5,247	3,518	1,729	694

establishments employing 20 workers or more maintain their gender distribution, while 7 and 6 per cent of the new and mature establishment, respectively, employing 10 or more workers, exhibit the same gender distribution at the two measurement times. Hence, small establishments are more likely to keep their gender distribution over time compared with larger ones.

The majority of new and mature establishments with 10 or more employees change their gender distribution by 10 percentage points or less between the two years observed. This also applies to new and mature establishment with 20 or more employees.

Very few establishments with 10 or more employees change their gender distribution by more than 25 percentage points, whereas 26 per cent of the new and 21 per cent of the mature establishments employing less than 10 workers do so. This big difference between small and larger establishments is due to the fact that in order to get a 25 percentage point change in its gender distribution an establishment with 4 workers, for example, only needs to replace one of the workers with a worker of the other sex, while an establishment with 100 employees has to replace 25 workers to get the same percentage point change.

A very small share of all establishments, regardless of size and age, change their gender distribution by more than 75 percentage points.

The change in gender distribution can be estimated in a more formal way by calculating the constant term in a first difference regression:

$$\text{gender distribution}_{i,8} - \text{gender distribution}_{i,2} = \alpha_8 - \alpha_2 + \varepsilon_{i,8} - \varepsilon_{i,2} \qquad (4)$$

The results are shown in Tables 6 and 7.[13] Positive coefficients show an increase in segregation and vice versa. Since we have a problem with regression towards the mean for establishment that are heavily gender segregated, we make separate analyses for establishments with different gender compositions. We also divide establishments by number of employees.

The main results are that all establishments become more integrated over time, except establishments with 50–74 per cent men which become more segregated. This is true for establishments of all sizes. New establishments employing 50–74 per cent men become more segregated, while new establishments employing 50–74 per cent women become more integrated than the corresponding mature establishments. The drivers behind this result are new, small establishments. It should be noted that the result that establishments employing more than 75 per cent men or 75 per cent women became less segregated could be a function of regression towards the mean.

Table 6. Changes in Gender Distribution in Male-Dominated Establishments between the First and Last Year of Observation.

	Mature Establishments: Share of Men in the First Year			New Establishments: Share of Men in the First Year		
	50–74%	75–89%	90–100%	50–74%	75–89%	90–100%
All	0.016*	−0.005*	−0.047*	0.037*	−0.020*	−0.084*
	0.002	*0.001*	*0.001*	*0.006*	*0.005*	*0.005*
	(12,259)	(11,471)	(6,693)	(1,215)	(1,194)	(911)
<10 employees	0.021**	−0.005**	−0.058**	0.048**	−0.018*	−0.084**
	0.003	*0.002*	*0.002*	*0.009*	*0.007*	*0.007*
	(6,511)	(6,661)	(3,891)	(758)	(814)	(643)
≥10 employees	0.011**	−0.005**	−0.031**	0.019**	−0.024**	−0.084**
	0.002	*0.001*	*0.001*	*0.006*	*0.006*	*0.008*
	(5,748)	(4,810)	(2,802)	(457)	(380)	(268)
≥20 employees	0.012**	−0.002	−0.023**	0.013	−0.022**	−0.066**
	0.002	*0.002*	*0.002*	*0.009*	*0.008*	*0.001*
	(3,184)	(2,523)	(1,351)	(207)	(171)	(90)

Notes: Standard errors are in italics. Number of establishments is in parenthesis.
**indicates significance at the 1 per cent level of confidence.
*indicates significance at the 5 per cent level of confidence.

Table 7. Changes in Gender Distribution in Female-Dominated Establishments between the First and Last Year of Observation.

Number of Employees	Mature Establishments: Share of Women in the First Year			New Establishments: Share of Women in the First Year		
	50–74%	75–89%	90–100%	50–74%	75–89%	90–100%
All	−0.024*	−0.050*	−0.056*	−0.048*	−0.093*	−0.083*
	0.002	*0.002*	*0.002*	*0.007*	*0.008*	*0.007*
	(10,645)	(7,072)	(3,529)	(1,180)	(718)	(514)
<10 employees	−0.031**	−0.058**	−0.060**	−0.060**	−0.104**	−0.083**
	0.003	*0.003*	*0.003*	*0.010*	*0.011*	*0.008*
	(6,037)	(4,648)	(2,622)	(710)	(513)	(421)
≥10 employees	−0.014**	−0.036**	−0.046**	−0.027**	−0.067**	−0.079**
	0.002	*0.002*	*0.003*	*0.008*	*0.010*	*0.013*
	(4,608)	(2,424)	(907)	(398)	(205)	(93)
≥20 employees	−0.017**	−0.032**	−0.041**	−0.028**	−0.084**	−0.098**
	0.002	*0.003*	*0.005*	*0.011*	*0.016*	*0.031*
	(2,283)	(1,053)	(305)	(144)	(80)	(24)

Notes: Standard errors are in italics. Number of establishments is in parenthesis.
**indicates significance at the 1 per cent level of confidence.
*indicates significance at the 5 per cent level of confidence.

3.4. Summary of Gender Distribution Dynamics

In conclusion, the Duncan and Duncan dissimilarity indexes and the Gini coefficients indicate that the overall segregation in the Swedish private sector has not changed between 1987 and 1995. According to these measures, the Swedish labour market is as segregated as the US manufacturing but less segregated than the Korean or Portuguese labour markets. Further, we find that new establishments are *not* more segregated than mature ones.

Plotting the establishments' gender distribution in the first and last years, we find a wide range of gender distributions in the individual establishments and in the way the distribution evolves over time. A majority of the establishments change their gender distribution within a 10 percentage point span. Almost one-fifth of them do not change their gender distribution at all during this period. From the first difference equations, we find that integrated firms with a small bias towards men (50–74 per cent) become slightly more segregated, and that all other establishments become slightly more integrated over time.

4. SURVIVAL

This section focuses on the relation between the gender distribution in an establishment and the probability of the establishment's survival. We start by estimating survivor functions using the Kaplan–Meier product-limit method to compare survival differences between female- and male-dominated establishments and establishments with a more even gender distribution.

An establishment is defined as *female dominated* if it belongs to the top 10 per cent of establishments with the highest ratio of the share of female employees relative to the share of female employees in the industry as a whole. An establishment is *male dominated* if it belongs to the 10 per cent of establishments with the highest ratio of the share of male workers in the establishment to the share of male workers in the industry as a whole. We use the ratio in order to control for the fact that some industries are heavily female or male dominated. All other establishments are defined as establishments with a more even (mixed) gender distribution.

The Kaplan–Meier survival estimates are shown in Fig. 2. Establishments are sorted by their gender distribution prevalent in year 2. According to the Kaplan–Meier survival estimates, female-dominated new establishments have significant higher survival than male-dominated new establishments and new establishments with a mixed gender distribution. There are no

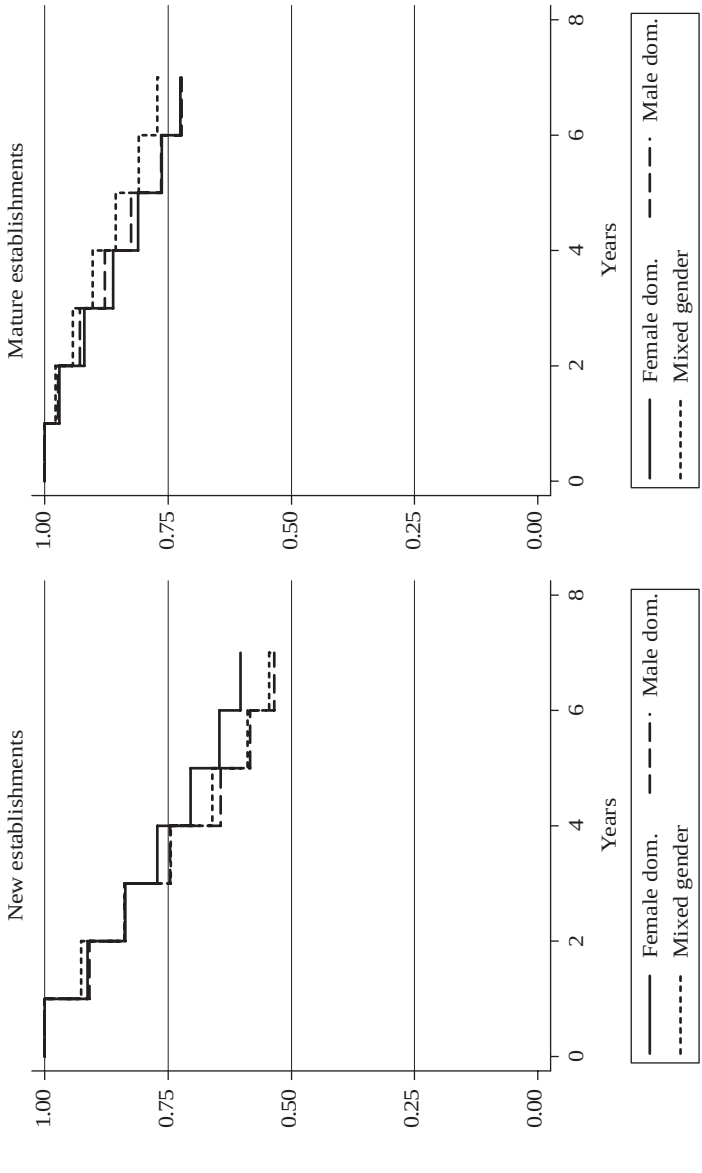

Fig. 2. Kaplan–Meier Survivor Function for Female- and Male-Dominated Establishments and Establishments with Mixed Gender Distribution.

significant differences in survival between male-dominated and mixed establishments.[14]

The results for the mature establishments differ from the results for the new establishments; mature establishments that are gender segregated have a significantly lower survival than more gender-integrated establishments. There is no statistical difference between male- and female-dominated establishments in terms of survival.

4.1. Method and Variables

In order to take more factors into account we estimate a discrete-time proportional hazard model with time-varying covariates since our data are interval-censored into years. That is, exact exit times are not known, only that they exit within some year. The baseline is non-parametric, i.e., we have created duration-interval-specific dummy variables, one for each year at risk, to estimate the baseline.[15] The presence of unobserved heterogeneity can produce misleading estimates of exit rates and attenuated estimates of covariates effects. One way to deal with this is to assume a specific distribution of the error term. In our data we find unobserved heterogeneity and therefore present models assuming normally distributed unobserved heterogeneity.[16]

Column 1 in Table 8 shows the risks, or the anti-logarithms, of estimated coefficients, of an exit for mature establishments. A value <1 implies that the factor reduces the risk of exiting. A value >1 implies that the factor has a positive effect on the risk of closing down. Column 2 shows the risk of an exit for new establishments.

In the models, we include various covariates that are expected to affect the risk that an establishment will exit. To examine the relation between gender composition and the risk to exit we construct time-varying dummy variables. First we include a dummy variable if the establishment is *female dominated* and another one if the establishment is *male dominated*. We use the same definitions of male- and female-dominated establishments as when estimating the Kaplan–Meier survivor functions. We repeat this procedure for all years. Hence, an establishment can be defined as female or male dominated in one year and gender mixed in another year.

The relationship among employees' age composition in a certain establishment, in a certain year and the establishment's probability of survival is examined as follows. We first sort the workers into four age groups: 16–24, 25–54, 55–59 and 60–64. We then define the top 10 per cent of establishments with the largest ratio of the per cent of workers in age

Table 8. Risk of Exiting.

Dependent Variable: Exiting	Hazard Ratio	
	Mature	New
Explanatory variables		
Female domination	1.629**	1.324**
	18.02	*5.30*
Male domination	1.657**	1.399**
	20.20	*7.05*
Dominated by workers aged 16–24	1.414**	1.116*
	11.59	*2.03*
Dominated by workers aged 25–54	1.754**	1.164**
	21.45	*3.24*
Dominated by workers aged 55–59	1.270**	1.049
	8.31	*0.83*
Dominated by workers aged 60–64	1.301**	0.961
	10.10	*−0.68*
Pre-upper secondary domination	1.414**	1.119*
	12.88	*2.13*
Secondary school domination	1.570**	1.306**
	16.90	*5.35*
University <3 years domination	1.128**	1.094
	3.93	*1.58*
University ≥ 3 years domination	1.240**	1.097
	6.70	*1.55*
Ln(employment in year 2)	0.826**	0.885**
	−13.32	*−3.84*
Employment growth since year 2, per year (%)	0.100**	0.113**
	−62.35	*−29.52*
Multi-unit	0.900**	0.922
	−4.30	*−1.54*
Dispersal		0.681**
		−8.29
Merger		0.761**
		−3.26
Birth year 1987		0.978
		−0.29
Controls for industries	Yes	Yes
No. of establishments	64,005	9,543
No. of closed establishments	15,378	4,296
No. of observations	349,241	45,470
No. of parameters	34	37
Wald χ^2	24,278.27	8,684.42
Log likelihood	−5,4053.844	−12,448.716

Notes: t-values are in italics.
**indicates significance at the 1 per cent level of confidence.
*indicates significance at the 5 per cent level of confidence.

group 16–24 in the establishment relative to the per cent of workers aged 16–24 in the industry as a whole. The variable thus defined is described as *dominated by age group 16–24*. We create a dummy variable equal to 1 if an establishment is dominated by age group 16–24 and 0 otherwise. We repeat the procedure for all years and for all age groups.

We also examine the relationship between the workers' educational level for each year and the establishments' survival probabilities. Workers are classified into four educational groups according to the highest education level they have achieved as follows: pre-upper secondary school, secondary school, less than three years of university education and three or more years of university education. We define the top 10 per cent of establishments as those with the largest per cent of workers with pre-upper secondary school relative to the per cent of such workers in the industry as a whole. The variable thus defined is defined as *dominated by pre-upper secondary school*. We create a dummy variable equal to 1 if an establishment is dominated by workers with education equal to pre-upper secondary school and 0 otherwise. We repeat this procedure for all years and for all educational groups.

Empirical studies suggest that the relation between size and the chances of survival can be expected to be positive, while the relation between size and employment growth can be expected to be negative (see, e.g., Persson, 1999). *Employment* measured as the logarithm of the number of employees and self-employed at the establishment in year 2 is therefore included as a control variable. In the regressions we include a time-varying variable on *average yearly employment growth since the second year*. The variable is defined as the logarithm of the number of employees in the establishment in the current year, less the logarithm of the number of employees in the establishment in the second year. This difference is then divided by the number of years since year 2.

We have also included variables that are not time varying, which are described below. We include a dummy variable, referred to as the 'multi-unit', which equals 1 if the establishment was a part of a larger firm with more than one establishment in the first measurement year.

A number of the new establishments have arisen as a result of mergers or dispersals, and these can be seen as 'artificial' births. Six per cent of the establishments are new due to mergers, and as many as 39 per cent are new due to dispersals. Establishments that are 'artificially' new may have better chances of surviving since they are already part of the market. We have therefore included one dummy variable that equals 1 if the establishment is new due to a *merger* and 0 otherwise, and one dummy variable that equals 1 if the establishment is new due to *dispersal* and 0 otherwise.

To control for possible effects of fluctuations in the business cycle on the survival probabilities of new establishments, we have included a dummy variable equal to 1 if the *birth year* was *1987* and 0 otherwise.

We also include *dummy variables for industries* since men and women are not evenly distributed between industries, and since establishments in some industries have a lower survival probability than those in other industries.

4.2. Results on Survival

The results from the discrete-time proportional hazard model with time-varying covariates are presented in Table 8. Female- and male-dominated establishments have a higher risk of exiting than establishments with mixed gender compositions. This result applies for both new and mature establishments. Hence, when controlling for different establishment and employee characteristics, the results from the Kaplan–Meier estimates suggesting that female-dominated new establishments have higher survival rates than male-dominated and gender mixed establishments are not longer true.

We have also elaborated with different cut-off points when defining if an establishment is male/female dominated. Besides the definition of female- and male-dominated establishments as the 10 per cent establishments employing the largest proportion of men and women, respectively, we have used 5 and 15 per cent as cut-off points. We find that the result that female- and male-dominated establishments have higher risk of exiting than other establishments are robust and not conditioned on what cut-off point we chose. In addition, we find that the more segregated the establishments are, the higher the risk of exiting.[17]

Further, we find that the 10 per cent of new establishments that employ the largest proportion of the youngest or the prime-age workers (aged 16–24 or 25–54, respectively) have a higher risk of exiting compared with establishments with a more even age distribution. For the mature establishments we find that establishments that belong to the 10 per cent establishments employing the largest proportion of any age group have higher exit risks than establishments with a more even age distribution.

Also, we find that particularly mature establishments with a more even educational distribution have lower exiting risks. The 10 per cent of mature establishments that employed the largest proportion of workers with any education have a higher risk of exiting than mature establishments with a more even educational distribution. New establishments belonging to the 10 per cent establishments employing the largest proportion of workers with

pre-secondary or secondary school have a higher risk than establishments with mixed educational distributions to exit the market.

The size of an establishment, in terms of number of employees, also affects its exit risk. The larger the establishment is in the second year, and the greater its average annual growth, the lower is the risk of exiting. New establishments due to dispersals or mergers have a lower risk of exiting than other new establishments in all periods.

Since the results for new and mature establishments are quite similar, we do not think that the left truncation of the data heavily bias the results for mature establishments.

5. GROWTH

In this section, we examine whether establishments employing a large proportion of women or men tend to expand as compared with establishments with a mixed gender distribution, i.e., if establishments dominated by one sex grow more in terms of employment relative to other establishments.

Table 9 shows regressions on employment growth between years 2 and 8, for establishments that survive until the eighth year. Establishment growth is defined as

$$\frac{\ln(\text{employment}_{t+7}) - \ln(\text{employment}_{t})}{7} \tag{5}$$

where t is equal to 1987 or 1988 depending on the establishment's birth year, and 1988 for mature establishments. The same explanatory variables as in Table 8 are used (except of course the yearly employment growth since the second year).

We find that both new and mature establishments dominated by females have a lower employment growth than new and mature establishments with a different gender composition.

Establishments that are dominated by workers from the youngest cohort or from the two oldest cohorts grow less than establishments with a more even age structure. The results are valid for both new and mature establishments.

The top 10 per cent of mature and new establishments employing the largest share of workers with a university education have a greater employment growth than establishments with less highly educated employees. In the previous section, we saw that establishments dominated by workers with at least three years of university education had a higher probability of exiting the market. From this we conclude that if establishments dominated

Table 9. Results from Regressions on Employment Growth OLS Estimates.

Dependent Variable: Average Annual Employment Growth				
Explanatory variables	Mature establishments		New establishments	
Female domination	−0.005**	*0.002*	−0.023**	*0.005*
Male domination	0.001	*0.001*	−0.005	*0.005*
Firm dominated by workers aged				
16–24	−0.007**	*0.002*	−0.016**	*0.005*
25–54	0.003	*0.002*	0.000	*0.005*
55–59	−0.011**	*0.002*	−0.012*	*0.005*
60–64	−0.018**	*0.001*	−0.017**	*0.005*
Firm dominated by workers with education from				
Pre-upper secondary school	−0.005**	*0.001*	−0.003	*0.005*
Secondary school	0.003	*0.002*	0.015**	*0.005*
University <3 years	0.007**	*0.002*	0.007	*0.006*
University ≥ 3 years	0.011**	*0.002*	0.029**	*0.006*
Ln(Employment in year 2)	−0.015**	*0.001*	−0.022**	*0.003*
Multi-unit	0.010**	*0.001*	0.006	*0.004*
Dispersal			−0.003	*0.004*
Merger			−0.005	*0.007*
Birth year 1987			−0.005	*0.003*
Controls for industries	Yes	Yes	Yes	Yes
Constant	0.004	*0.002*	0.048	*0.008*
No. of observations	48,627		5,247	
Root MSE	0.010		0.119	

[a]Using Cook–Weisberg tests for heteroskedasticity indicates that standard errors are heteroskedastic. We have therefore run the regressions with robust standard errors. Standard errors are in italics.
**indicates significance at the 1 per cent level of confidence.
*indicates significance at the 5 per cent level of confidence.

by workers with at least three years of university education survive they are likely to inherit a substantial growth.

6. CONCLUSIONS

We study gender segregation in privately owned Swedish establishments, and the correlation between gender segregation and the survival and growth of establishments. Different theoretical models give different implications

about gender segregation, employment dynamics and firm profitability. Becker's (1957) model on employer discrimination predicts that firms that employ a large fraction of women will be relatively more profitable due to lower wage costs, and thus enjoy a greater probability of growing by underselling other firms in the competitive product market.

Like Becker's 'taste'-based model of discrimination, Lang's (1986) language model implies firm segregation in the long run. Mello and Ruckes (2006) present a model of team composition where heterogeneous teams have greater variety of information sources than homogenous teams. If information and preferences can be expressed openly, heterogeneous work teams reach better decisions. In the corporate governance literature, gender diversity and firm performance has been studied to a quite large extent and both the theoretical and empirical evidence are ambiguous (see Smith et al., 2006).

We find that the Swedish private sector is as segregated as the US manufacturing, but less segregated than the Portuguese or Korean labour markets. Over time, establishments with 50–75 per cent male workers become more segregated (male dominated) while all other establishments become more integrated. These results speak against models that predict gender segregation in the long run, as for example, Lang's and Becker's models.

We also find that establishments dominated by males or females have a higher probability of exiting the market than more integrated establishments. However, establishments dominated by females grow more slowly than other establishments. This holds for both new and mature establishments. These results suggest that the predictions from Becker's model about female-dominated establishments' higher survival and growth rates, compared to male-dominated or mixed establishments, are not valid for Sweden. Instead, theories suggesting that workers with different demographic characteristics contribute to a creative working environment as a result of their different experiences, a greater variety of information sources and different 'thinking' (e.g., Mello & Ruckes, 2006) are hence supported by the Swedish data.

An important additional finding is that establishments with a skewed workforce in terms of educational background have lower survival probabilities. Kremer and Maskin (1996) have developed a model on wage inequality and segregation by skill. Their model predicts that countries with greater educational dispersion have firms that are more segregated by education than countries with lower educational dispersion. Hence, if a country has a more compressed educational distribution as is the case in

Sweden, firms should be more heterogeneous. This prediction is in line with our empirical result on Swedish privately owned establishments.

Furthermore, establishments with skewed age distributions have lower survival probabilities and also grow less compared with other establishments.

To conclude, integrated, heterogeneous establishments seem to be more successful than other establishments in Sweden. Thus, attempts by legislators to integrate firms along all dimensions of *diversity* may have positive effects on the growth and survival of firms.

NOTES

1. In the US for example, employers with federal contracts and 50 or more employees, or with contracts worth more than $50,000, have to file reports on skewed gender distribution and affirmative actions (Executive Order 11246). According to European Union law on equal treatment (Article 141 and Directive 76/207/EEC), all gender segregation in the labour market is prohibited and the European Council recommends member states to adopt affirmative action to remove gender inequalities (Directive 84/635/EEC). According to Swedish law, if the gender distribution is unequal in a workplace, the employer has to endeavour to recruit applicants of the underrepresented sex, in order to gradually increase its representation. Firms with more than 10 employees are also obliged to prepare a plan of action aiming at gender equality in the workplace (SFS 1991:433, replacing SFS 1979:1118).

2. Groshen (1991) examines US industries and finds that wages are negatively correlated to the proportion of women in an establishment, and that about 6 per cent of the gender wage gap could be explained by establishment segregation. Bayard, Hellerstein, Neumark, and Troske (2003) study wages and segregation in all sectors in the US in 1990. They find that 16–17 per cent of the wage gap is attributable to establishment segregation. Analysing the Swedish private sector, Arai, Nekby, and Thoursie (2004) find that establishment gender segregation explains 7 per cent of the gender wage gap. Carrington and Troske (1998) find that inter-plant gender segregation in the US manufacturing in 1990 was substantial, and that inter-plant gender segregation could account for a substantial fraction of the male/female wage gap. They also found that men who work in female-dominated establishments had lower wages, on average, than other men. In a study by Carrington and Troske (1995) inter-firm gender segregation was found to be prevalent among small US employers, and sex segregation accounted for a large part of the gender wage gap. Cabral Vieria, Cardoso, and Portela (2005) analyse gender segregation across Portuguese establishments and find it to be high and stable over time. They find that the higher the concentration of women in the establishment, the lower the women's wages. They also find, in contrast to Carrington and Troske (1998), that men working in female-dominated establishments receive higher wages than other men. The gender segregation and its impact on wages in Korea is examined by Yoon, Troske, and Mueser (2003), who find that segregation across establishments plays an important role in explaining the gender wage gap. Black and Brainerd (2004) test

Becker's hypothesis that increased market competition will drive out costly discrimination in the long run. They find that when competition increases as a function of trade, the gender wage gap decreases, which supports Becker's predictions.

3. Women's lower wages can also be explained by differences in human capital investment as suggested by Mincer and Polachek (1974). According to them, the behavior of the family implies a division of labour within it. The family's differential allocation of time and investments in human capital is generally sex linked.

4. There are few empirical studies on the relationship between profitability and discrimination. Some of them are discussed in Rosén (2003).

5. An establishment is defined as an address (not a household address), a building or a group of adjacent buildings where a firm operates. When it is a question of a mobile activity (e.g., home-help services), a succession of temporary work sites, work spread over a large area or work consisting of renting out premises or apartments, then the activities of the firm are assigned to the location from which the work is administered. If the firm has one establishment only, that establishment is taken as synonymous with the firm.

6. The construction industry, which accounts for around 6 per cent of total employment, is excluded. The reason for this is that establishments in the construction industry are mobile and connected to building sites, which makes it difficult to define new or existing establishments in a meaningful way.

7. See Persson (1999) for a more detailed description of the data and how it was compiled.

8. Sometimes the establishment is inactive for a particular year, or the information is missing. If an establishment is inactive for one year or more, i.e., to say no individual is connected to it, then there will be no information about the establishment for that year (or those years). The lack of information for a particular year is sometimes due to the absence of the figures concerned, but according to Statistics Sweden most of these missing years in fact reflect the inactivity of the establishment. We exclude all establishments that lack information for one year or more.

9. An establishment is considered as an entry in year t if its identity number has been assigned during that year and does not occur during previous years (1985 and 1986 for those that are new in 1987, and 1985–87 for those that are new 1988). In the analyses, we study the establishments from the second year of their existence, i.e., from 1988 to 1994 for establishments created in 1987 and from 1989 to 1995 for establishments created in 1988. This is because a large share of new establishments already exits during the first year.

10. We have also made sensitivity analyses using sub-samples with establishments employing at least 6 employees and at least 10 employees in the second year. The results are almost identical. However, only 50 per cent of the new establishments employing at least 6 employees and 25 per cent of the new establishments employing at least 10 employees survived the first year. The corresponding figures for mature establishments are 62 respectively 40 per cent.

11. See Carrington and Troske (1998).

12. To calculate G^* we first randomly reallocated the workers between the firms. After the reallocation, the Gini coefficient was computed. This was repeated 50 times. We then calculated the average Gini coefficient and defined it as G^*.

Cabral Viera et al. (2005) use the same technique. The D^* is calculated analogously. Averages, standard deviations, and minimum and maximum values for G^* and D^* are presented in Tables A1 and A2. A special thanks here to Jan Selén who helped us with the computer programming.

13. For coefficients for separate years, see Tables A3 and A4.

14. According to both log-rank (or Savage) and Wilcoxon–Breslow tests, the differences between the survivor functions are statistically significant at the 1 per cent level. Both tests are used, since the Wilcoxon–Breslow test stresses differences in the survivor function at the beginning of the duration while the log-rank test stresses increasing differences at the end of the process time (see, e.g., Blossfeld, Golsch, & Rohwer, 2007).

15. For a description of the discrete-time proportional hazard model (sometimes referred to as the cloglog model due to the complementary log–log transformation), see Jenkins (2005).

16. We have also estimated the models assuming gamma-distributed unobserved heterogeneity. This did not change our results.

17. The magnitude of the coefficient is larger for the 5 per cent of establishments that employ the largest proportion of women/men than for the 10 per cent, which in turn is larger than the coefficient for the 15 per cent of the establishments that employ the largest fraction of men/women in the industry.

ACKNOWLEDGMENTS

We have benefited from comments on various versions of this chapter during seminars at the following institutes: Swedish Institute for Social Research, the Department of Economics at Stockholm University, the EALE conference in Jyväskylä, the ESPE conference in Athens and the Workshop on the Economic Analysis of Linked Employer-Employee Data in Århus; all of which we gratefully acknowledge. We would also like to thank Gerard van den Berg, Maria Hemström, Matthew Lindquist, Åsa Rosén, Åsa Segendorf, Jan Selén, Konstantinos Tatsiramos, Eskil Wadensjö and two anonymous referees for helpful suggestions. Helena Persson would also like to thank the Bank of Sweden Tercenterary Foundation for financial support.

REFERENCES

Arai, M., Nekby, L., & Thoursie, P. S. (2004). *Is it what you do or where you work that matters most? Gender composition and the gender wage gap revisited.* Working Papers in Economics no. 2004:10. Department of Economics, Stockholm University, Stockholm.

Bayard, K., Hellerstein, J., Neumark, D., & Troske, K. (2003). New evidence on sex segregation and sex differences in wages from matched employee-employer data. *Journal of Labor Economics*, *21*(4), 887–922.

Becker, G. S. (1957). *The economics of discrimination*. Chicago, IL: University of Chicago Press.

Black, S. E., & Brainerd, E. (2004). Importing Equality? The impact of globalization on gender discrimination. *Industrial and Labor Relations Review*, *57*(4), 540–559.

Blossfeld, H.-P., Golsch, K., & Rohwer, G. (2007). *Event history analysis with stata*. New York: Lawrence Erlbaum Associates, Inc.

Cabral Vieria, J. A., Cardoso, A. R., & Portela, M. (2005). Gender segregation and the wage gap in Portugal: An analysis at the establishment level. *Journal of Economic Inequality*, *3*(2), 145–168.

Carrington, W. J., & Troske, K. R. (1995). Gender segregation in small firms. *Journal of Human Resources*, *30*(3), 503–533.

Carrington, W. J., & Troske, K. R. (1997). On measuring segregation in samples with small units. *Journal of Business and Economic Statistics*, *15*(4), 402–409.

Carrington, W. J., & Troske, K. R. (1998). Sex segregation in U.S. manufacturing. *Industrial and Labor Relations Review*, *51*(3), 445–464.

Groshen, E. L. (1991). The structure of the female/male wage differential. *Journal of Human Resources*, *26*(3), 457–472.

Haltiwanger, J. C., Lane, J. I., & Spletzer, J. R. (2007). Wages, productivity, and the dynamic interaction of businesses and workers. *Labour Economics*, *14*(3), 575–602.

Hellerstein, J. K., Neumark, D., & Troske, K. R. (2002). Market forces and sex discrimination. *Journal of Human Resources*, *37*(2), 353–380.

Jenkins, S. P. (2005). *Survival analysis*, unpublished manuscript. Available at www.iser.essex.ac.uk/files/teaching/stephenj/ec968/pdfs/ec968lnotesv6.pdf

Kremer, M., & Maskin, E. (1996). *Wage inequality and segregation by skill*. NBER WP 5718.

Lang, K. (1986). A language theory of discrimination. *Quarterly Journal of Economics*, *101*(2), 363–382.

Mello, A. S., & Ruckes, M. E. (2006). Team composition. *Journal of Business*, *79*, 1019–1039.

Mincer, J., & Polachek, S. (1974). Family investments in human capital: Earnings of women. *The Journal of Political Economy*, *82*(2, part 2), s76–s108.

Persson, H. (1999). *Essays on Labour Demand and Career Mobility*. Dissertation series no. 40. Swedish Institute for Social Research, Stockholm University.

Rosén, Å. (2003). Search, bargaining and employer discrimination. *Journal of Labor Economics*, *21*(4), 807–829.

Smith, N., Smith, V., & Verner, M. (2006). Do women in top management affect firm performance? A panel study of 2,500 Danish firms. *International Journal of Productivity and Performance Management*, *55*(7), 569–593.

Yoon, S., Troske, K. R., & Mueser, P. (2003). *Changes in gender segregation and women's wages in Korea*. Mimeo, University of Missouri-Colombia.

APPENDIX

Table A1. Random Gender Segregation, Mature Establishments.

Year	The Random Dissimilarity Index				The Random Gini Coefficient			
	D^*	SD	Minimum value	Maximum value	G^*	SD	Minimum value	Maximum value
2	0.127	*0.000*	0.126	0.128	0.201	*0.000*	0.200	0.203
3	0.127	*0.004*	0.125	0.153	0.200	*0.001*	0.199	0.201
4	0.127	*0.000*	0.126	0.128	0.201	*0.001*	0.200	0.203
5	0.129	*0.001*	0.128	0.130	0.204	*0.001*	0.202	0.205
6	0.133	*0.001*	0.132	0.139	0.210	*0.001*	0.209	0.211
7	0.136	*0.001*	0.134	0.137	0.215	*0.001*	0.212	0.216
8	0.132	*0.001*	0.131	0.133	0.209	*0.001*	0.207	0.211
No. of establishments in year 2	64,005	64,005	64,005	64,005	64,005	64,005	64,005	64,005
No. of establishments in year 8	48,627	48,627	48,627	48,627	48,627	48,627	48,627	48,627

Note: Standard deviations (SD) are in italics.

Table A2. Random Gender Segregation, New Establishments.

Year	The Random Dissimilarity Index				The Random Gini Coefficient			
	D^*	SD	Minimum value	Maximum value	G^*	SD	Minimum value	Maximum value
2	0.208	*0.002*	0.203	0.216	0.314	*0.003*	0.309	0.319
3	0.204	*0.002*	0.201	0.209	0.311	*0.003*	0.307	0.318
4	0.203	*0.002*	0.199	0.207	0.309	*0.002*	0.304	0.314
5	0.204	*0.002*	0.201	0.209	0.311	*0.003*	0.305	0.318
6	0.204	*0.002*	0.201	0.209	0.313	*0.003*	0.307	0.319
7	0.199	*0.002*	0.194	0.203	0.304	*0.003*	0.299	0.310
8	0.190	*0.002*	0.186	0.194	0.292	*0.003*	0.287	0.297
No. of establishments in year 2	9,543	9,543	9,543	9,543	9,543	9,543	9,543	9,543
No. of establishments in year 8	5,247	5,247	5,247	5,247	5,247	5,247	5,247	5,247

Note: Standard deviations (SD) are in italics.

Table A3. Changes in Gender Distribution in Male-Dominated Establishments.

Year	Mature Establishments: Share of Men in the First Year			New Establishments: Share of Men in the First Year		
	50–74%	75–89%	90–100%	50–74%	75–89%	90–100%
2–3	0.000	−0.010**	−0.026**	0.005	−0.019**	−0.046**
	0.001	*0.001*	*0.001*	*0.004*	*0.004*	*0.003*
3–4	−0.001	−0.004**	−0.010**	0.002	−0.002	−0.012**
	0.001	*0.001*	*0.001*	*0.004*	*0.004*	*0.003*
4–5	0.003**	−0.004**	−0.005**	0.005	−0.001	−0.009**
	0.001	*0.001*	*0.001*	*0.004*	*0.004*	*0.003*
5–6	0.006**	0.002*	−0.001	0.013**	−0.002	−0.007*
	0.001	*0.001*	*0.001*	*0.004*	*0.004*	*0.004*
6–7	0.010*	0.002*	−0.002*	0.010*	−0.001	−0.001
	0.004	*0.001*	*0.001*	*0.004*	*0.004*	*0.003*
7–8	0.001	0.001	−0.003*	0.003	0.006	−0.009*
	0.001	*0.001*	*0.001*	*0.004*	*0.004*	*0.004*
2–8	0.016**	−0.005**	−0.047**	0.037**	−0.020**	−0.084**
	0.002	*0.001*	*0.001*	*0.006*	*0.005*	*0.005*
No. of observations	12,259	11,471	6,693	1,215	1,194	911

Notes: Standard errors are in italics.
**indicates significance at the 1 per cent level of confidence.
*indicates significance at the 5 per cent level of confidence.

Table A4. Changes in Gender Distribution in Female-Dominated Establishments.

Year	Mature Establishments: Share of Women in the First Year			New Establishments: Share of women in the First Year		
	50–74%	75–89%	90–100%	50–74%	75–89%	90–100%
2–3	−0.006**	−0.019**	−0.026**	−0.013**	−0.044**	−0.047**
	0.001	*0.001*	*0.001*	*0.005*	*0.006*	*0.004*
3–4	−0.003*	−0.006**	−0.007**	−0.003	−0.006	−0.019**
	0.001	*0.001*	*0.001*	*0.004*	*0.006*	*0.004*
4–5	−0.003**	−0.008**	−0.004**	−0.006	−0.021**	−0.008
	0.001	*0.001*	*0.001*	*0.004*	*0.005*	*0.005*
5–6	−0.007**	−0.007**	−0.010**	−0.013**	−0.017**	−0.004
	0.001	*0.001*	*0.001*	*0.005*	*0.006*	*0.005*
6–7	0.002	−0.001	−0.003*	−0.006	−0.009	−0.004
	0.001	*0.001*	*0.001*	*0.005*	*0.005*	*0.006*
7–8	−0.006**	−0.010**	−0.006**	−0.006	0.004	−0.001
	0.001	*0.002*	*0.001*	*0.005*	*0.006*	*0.005*
2–8	−0.024**	−0.050**	−0.056**	−0.048**	−0.093**	−0.083**
	0.002	*0.002*	*0.002*	*0.007*	*0.008*	*0.007*
No. of observations.	10,645	7,072	3,529	1,180	718	514

Notes: Standard errors are in italics.
**indicates significance at the 1 per cent level of confidence.
*indicates significance at the 5 per cent level of confidence.

FUTILE AND EFFECTIVE WAYS TO COMBAT WAGE DISCRIMINATION

Yuval Shilony and Yossef Tobol

ABSTRACT

Using Becker's 'taste for discrimination' model, the chapter analyzes the current legislation against wage discrimination and finds it counter-productive. Using a costly apparatus of auditing, detecting and fining violators does not deliver results. If a fine is levied on discriminators and reimbursed to the disadvantaged workers in order to undo the discrimination, it affects equally the demand for and the supply of those workers, because their expected wage includes the fine, and has no real effect. If the fine is collected and kept by the government, it shifts employment away from the workers it seeks to help, to others, depressing the total employment. In contrast, levying a tax on the favored workers effectively curbs discrimination in the labor market. A quota is a possible substitute for a tax with questionable side effects. Affirmative action is in essence a sort of tax on employing favored workers, only administered in an indirect, clumsy and costly way. Yet, the chapter explains its humble impact in the right direction. An explicit and direct tax would do much more and with a negative cost. Alternatively, subsidizing the disfavored workers is a costly but as effective policy that, in addition, boosts total employment.

Jobs, Training and Worker Well-Being
Research in Labor Economics, Volume 30, 283–300
Copyright © 2010 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 0147-9121/doi:10.1108/S0147-9121(2010)0000030012

1. INTRODUCTION

Wage discrimination is vastly studied by economic theory beginning with Becker (1957, 1971) and extended by Arrow (1972a, 1972b, 1974). According to Becker (1971), an individual has a taste for discrimination if he is willing to pay to be associated with some persons rather than others. True, in competitive markets nondiscriminators have an advantage but discriminators may be willing to forgo profit and income to indulge in their taste and survive the competition.

In the United States, employment discrimination is restrained by a set of five laws administered by the Equal Employment Opportunity Commission (EEOC):

1. The Equal Pay Act of 1963, prohibiting gender-based wage discrimination.
2. Title VII of the Civil Rights Act of 1964, prohibiting employment discrimination based on race, color, religion, sex and national origin.
3. The Age Discrimination in Employment Act of 1967 (ADEA), prohibiting employment discrimination against persons 40 years of age or older.
4. The Americans with Disabilities Act of 1990 (ADA), prohibiting employment discrimination against qualified individuals with disabilities.
5. The Civil Rights Act of 1991 modifying the prior laws to allow compensatory and punitive damages in cases of employer 'malice or reckless indifference to the rights of aggrieved individuals'.

Empirical work has tended to show that the Civil Rights laws and EEOC enforcement improve economic conditions for protected groups, (see, e.g. Heckman & Payner, 1989; Leonard, 1996). Other works were concerned that the laws actually induce discrimination (see Posner, 1987; Donohue & Siegelman, 1991; Abram, 1993; Acemoglu & Angrist, 2001; Oyer & Schaefer, 2000, 2002). For example, Oyer and Schaefer (2002) find that employers avoid hiring blacks and women in order to forestall later employment litigation.

The literature on discrimination as a crime and the implications for firm's management and employment is rather short. Tobol (2005) examined the implications of enforcing the Equal Pay Law in monopsonistic and competitive markets. Other than that, the closest economic theory is the minimum wage law and its noncompliance (see, e.g., Chang & Ehrlich, 1985; Yaniv, 2001).

This chapter evaluates the effectiveness of policies, not the merit of the target. In the background there is a social decision to curb discrimination,

deemed socially bad. First, two alternative laws levying a fine on discriminators, are examined, both futile: (1) the fine is collected and kept by the government and (2) the fine is reimbursed to the disfavored workers, thus feeding back on labor supply. The framework is laid out in Section 2 of this chapter and the results stated and proved in Section 3. In contrast, there are two alternative ways, both expedient means in curbing wage discrimination without the need to monitor, catch and punish discriminators. The first is a fine on employing favored workers, dealt with in Section 4. Note that affirmative action[1] is in essence a tax on employing A-type workers, as modeled and empirically assessed by Leonard (1996), only administered in an indirect, clumsy and costly way. Still, the result in Section 4 explains its humble impact in the right direction. An explicit and direct tax would do much more and with negative costs. The second is a subsidy for employing disfavored workers and is analyzed in Section 5. In Section 6, a quota is discussed as a substitute for a tax. Section 7 holds concluding remarks.

2. ANALYSIS OF CURRENT APPROACH AND LEGISLATION

Consider an employer producing a given product, whose price is 1, in a competitive market with two types of workers, A and B (e.g., males and females, whites and blacks), where A and B denote also the quantity employed of each type. Denote the wage rates of A- and B-workers by w_A and w_B, respectively. To focus on discrimination, suppose A- and B-workers have identical productive skills, hence the employer's production function is given by $f(A+B)$. The marginal product is positive and diminishing, i.e., $f'>0, f''<0$. Profit is given by

$$\pi = f(A+B) - w_A A - w_B B \tag{1}$$

If the wages of the two types of workers differ, any profit-maximizing employer would hire only the cheaper workers. Therefore equilibrium in the labor market implies equal pay and no discrimination.

Recognizing this, Becker (1957) and Arrow (1972a, 1972b), replaced profit maximization with *utility* maximization. A wage differential stems from a taste for discrimination against B-workers in spite of their identical productive skills. Specifically, the employer's utility function is assumed to be $U = U(\pi, A, B)$, where $U_\pi>0$, $U_A>0$ but $U_B<0$. To simplify, assume the

utility function is additively separable in the three arguments:

$$U(\pi, A, B) = \pi + \psi(A) - \phi(B) \tag{2}$$

where $\psi'(A)>0$, $\phi'(B)>0$ and $\psi''(A)<0$, $\phi''(B)>0$ and profit is given by Eq. (1).

However, if wage discrimination is illegal, a risk is introduced. Following the literature on minimum wage noncompliance (e.g., Chang & Ehrlich, 1985; Yaniv, 2001), we assume that if discrimination is practiced and caught, the employer is fined in proportion to the wage underpayments. That is, the fine will be some proportion $\lambda(>1)$ of the unpaid wages to B-workers, $[wA-wB]B$.

With the law in effect expected profit is

$$E\pi = f(A+B) - w_A A - w_B B - p\lambda(w_A - w_B)B \tag{3}$$

where p denotes the probability of getting caught and punished, assumed to be independent of the wage differential or the employment levels.

Assume there is a continuum of price-taker producers-employers indexed by t, $t \in [0,1]$. They are identical in every respect except for preferences. Employer t employs $A(t)$ and $B(t)$ and has preferences $\psi(t, A(t))$ and $\phi(t, B(t))$. Suppose t indicates the intensity of the employer's preference for A-workers and dislike of B-workers. So for $t_1<t_2$ and any A and B, $\psi'(t_1, A)<\psi'(t_2, A)$ and $\phi'(t_1, B)<\phi'(t_2, B)$, where $\psi'(t, A) = (\partial\psi(t, A))/\partial A$ and $\phi'(t, B) = (\partial\psi(t, B))/\partial B$. The types of employers t are distributed according to density $g(t)$. Maximizing Eq. (2), employers with a large enough t would choose not to employ B-workers at all, adopting a corner solution. Note that, rewriting employer t's target, denoting $e = p\lambda$ and assuming $e<1$, one gets:

$$\underset{A,B}{\text{Max}}\ EU = \{f(A+B) - w_A A - w_B B - e(w_A - w_B)B + \psi(t, A) - \phi(t, B)\}$$
$$\textit{subject to constraint } B \geq 0 \tag{4}$$

The first-order conditions for a maximum are

$$\frac{\partial(EU)}{\partial A} = f'(A+B) - w_A + \psi'(t, A) = 0 \tag{5}$$

$$\frac{\partial(EU)}{\partial B} = f'(A+B) - w_B - e(w_A - w_B) + \phi'(t, B) = 0 \tag{6}$$

implying equality of the marginal revenue from employment, $f'(A+B)$, with the expected marginal cost. Notice that the marginal monetary cost of

employing A-workers is moderated by their marginal utility to the employer, $\psi'(A)$, whereas the marginal cost of employing B-workers is augmented by their marginal disutility to the employer, $\phi'(B)$. Solving Eqs. (5) and (6), one gets t's demands for workers $A(t, w_A, w_B, e)$ and $B(t, w_A, w_B, e)$.

First note that for a large enough t^*, Eq. (6) vanish for $B = 0$; so all employers t with $t^* < t \leq 1$ dislike B-workers so much that Eq. (6) turns negative. Hence they optimally choose $B(t) = 0$. All other employers use both types of workers. Another variety, not followed here, of the anti-B preference may be $\phi''(B) < 0$, i.e., declining marginal displeasure from Bs, implying something like 'one bad apple spoils the barrel' idea. It burdens the optimization since Eq. (6) may then lead to a local minimum, and increase the proportion of employers who choose a corner solution with no Bs.

Note that by avoiding B-workers the employer may avert the appearance of wage discrimination. One does not employ any B-workers, does not pay them lower wages than A-workers and does not exploit the B-workers. In fact, it is a more severe form of discrimination, discrimination in hiring. The B-workers suffer twice: wage discrimination and hiring discrimination, implying suppressed demand for their services.

Let the supply functions of the two types of workers be $S_A(w_A)$ and $S_B(w_B)$. Labor market equilibrium implies

$$\begin{aligned} S_A(w_A) = \bar{A}(w_A, w_B, e) = \int_0^1 A(t, w_A, w_B, e) g(t) \mathrm{d}t \text{ and} \\ S_B(w_B) = \bar{B}(w_A, w_B, e) = \int_0^1 B(t, w_A, w_B, e) g(t) \mathrm{d}t \end{aligned} \tag{7}$$

where $\bar{A}$ and $\bar{B}$ are total market demands.

Market equilibrium is attained at the solution of this system of equations: Eqs. (5) and (6) for each producer t and the two ones of the aggregate labor market Eq. (7). For each employer t who uses both types, we can deduce from Eqs. (5) and (6) that by our assumptions the wage levels differ:

$$w_A - w_B = \frac{\psi'(t, A) + \phi'(t, B)}{1 - e} > 0 \tag{8}$$

A crucial issue is the use entailed by law for the fine's proceeds. The money can simply go to government's coffer. Call this policy Law No. 1. It was analyzed by Tobol (2005) for the case of vertical labor supplies. Claim 1 generalizes the analysis of that policy in several respects and is instrumental for the further results. Another law may attempt to undo the wrong and pay the proceeds of the fine to the aggrieved B-workers of the fined employer. Call this policy Law No. 2. It also affects the supply of the B-workers and is tackled in Claim 2.

3. RESULTS

We examine several policies designed to curtail wage discrimination in the labor market. The policy tools are legislated taxes and subsidies. Wage discrimination is viewed as an illegal activity and the law is enforced by assumption. We examine how each suggested policy influences the wages, employment, employees and employers. In addition to the two laws mentioned (the fine revenue kept by the government or passed to the disadvantaged workers), we consider levying tax on the advantaged workers or giving subsidy to the disadvantaged workers. We show that these two latter are not symmetric.

Claim 1. Assume a competitive labor market as described above, with two types of productively identical workers, A and B, having rising labor supply curves. Employers have a discriminating preference against B-workers and the law levies a fine amounting to the difference in wage times the number of the employer's disadvantaged workers $(w_A - w_B)B$. Assume the B-labor demand depends on the expected wage $ew_A + (1 - e)w_B$. The revenue from the fine is kept by the government. Making enforcement of the law, e, stronger, is counterproductive by:

1. Depressing total employment.
2. Increasing the wage and employment of A-workers and decreasing the wage and employment of B-workers.
3. Expanding the set of employers who avoid B-workers altogether.

The proof is in the appendix.

Intuition and discussion:
The probability of catching any discriminating employer is e. If caught, the employer pays the B-workers wage w_A. Otherwise, with the probability $1-e$, he pays only w_B. Therefore, the expected wage is $ew_A + (1 - e)w_B$. From Eq. (4) one sees that e does not change the incentive to employ A-workers. On the other hand it raises the (random) cost of employing B-workers. So, all in all, the cost of employment increases, thus depressing total employment. Of course, the B-workers are hurt more because their employment is made more costly not only in relative terms, i.e., their relative wage to the employer goes up, but even in absolute terms, implying a tendency to employ less of them, even avoiding them altogether by more firms.

The discriminating employer anticipates the expected fine and considers it an additional cost. He takes optimal measures to compensate for that and defend himself in advance by reducing his demand for *B*-workers and increasing that for *A*-workers. Of course his behavior and willingness to pay still higher wages to *A*-workers stems from his utility from employing *A*-workers and disutility from employing *B*-workers. From a micro perspective, Law No. 1 is counterproductive, even illogical, in a competitive market. The government hurts the workers it aims to help. Note that Tobol (2005) has shown that the very same policy would achieve its goal in a monopsonistic market. However, Claim 1 provides a clue for a useful direction. If this policy is counterproductive – do the opposite one, and move in the reverse direction!

Now examine the tax policy when the revenue is passed to the disadvantaged workers instead of increasing the government budget surplus. While under Law No. 1 only the employers faced uncertainty regarding the cost of *B*-labor, i.e., the actual *B*-wage paid, now both employers and *B*-workers share it.

Claim 2. In the competitive labor market as described above, assume the fine, if and when levied on any employer, is paid to his *B*-workers. Assume that in addition to demand, the *B*-labor supply is now dependent on the expected wage, $ew_A + (1-e)w_B$, as well. Enforcement of Law No. 2 at any strength, i.e. $e \geq 0$, has no real effect. Its only effect is to adjust the market wage of *B*-workers so as to keep their expected wage constant. Both employment levels remain intact. Law No. 2 is equivalent in real terms to having no law at all.

The proof is in the appendix.

Intuition and discussion:
Both the demand (of each firm and thus the total one) for, and the supply of, *B*-workers now depend on the expected wage. Any change of it, due to variation of e, affects the demand price and the supply price equally. So in the *B*-labor market both curves shift down equally when e is increased.

Fig. 1 demonstrates the *B*-labor market. The solid curves are in terms of the expected wage, the effective wage that directs the decisions of both workers and employers, by our behavioral assumption, and intersect in point *E*. However, the wage, $w_B^* = ew_A + (1-e)w_B$, is not transacted in the market, it is the expectation of a lottery taken by the employers who wage-discriminate. The wage actually paid to each *B*-worker w_B is lower as it is net of any fine.

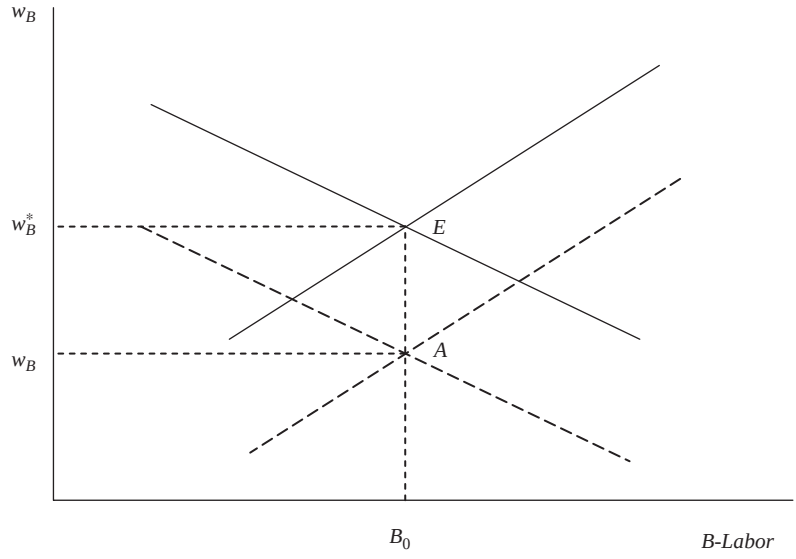

Fig. 1. The *B*-Labor Market: Transacted vs. Expected Wage.

The dashed curves are in terms of the transacted wage, the wage actually paid by the employer to the worker, and intersect in point A. The fine, if and when applicable, is paid by the government, not by the employer. The difference between the two levels of wage, $w_B^* - w_A = e(w_A - w_B)$, depends on the policy variable e. Increasing e would shift the dashed curves in the figure further down, only reducing the transacted wage but not affecting any real magnitude. It is like subsidizing the seller and taxing the buyer of a product to the same extent per unit. The two cancel out each other. Employment does not change. Wage only seems to go down. In addition to the lower transacted market wage, employers pay the tax and workers receive the transfer, all to restore the old higher wage. Equilibrium is the same in both *A*- and *B*-labor markets, so the law/policy has zero effect.

A Note on Risk Aversion

Our assumption that the demand for and supply of *B*-workers depend on the expected wage, and not on any other aspect of the distribution of the risk taken, amounts to assuming risk neutrality on the part of workers and employers. If, however, participants are, as usual, risk averse, it might

change their behavior and decisions. It is not difficult to anticipate the direction of change. For a risk-averse producer, facing uncertain input price, the risk adds a burden to the marginal factor cost that is tantamount to a rise in the input price, which has two effects. It would direct the producer to economize on the use of the input and use substitutes instead. Since production is more expensive, compared to a certain input price, it negatively affects the quantity produced. Both effects spell bad for the labor market and the policy. In Claim 1, it implies more pronounced effects of the tax, less B-employment and more of A. In the case of Claim 2, the neutrality of the tax-cum-subsidy is gone. The employers' demand for B-labor is reduced because of the risk burden introduced. The workers' supply is reduced for the same reason. Together they imply less B-employment and more of A. The following alternative measures do not introduce risk and so are not affected by the risk-neutrality assumption.

4. EFFECTIVE ALTERNATIVE POLICIES: TAX

Now examine a policy of taxing of the favored workers. This policy is easier to enforce, as it does not discriminate between employers who wage-discriminate and who do not, and it does not require detection and catching. Of course, the tax revenue is kept by the government and boosts the government's budget.

Suppose a tax of K is levied on each A-worker employed. Now employer t's target is

$$\begin{gathered}\underset{A,B}{\text{Max}}\, U = f(A+B) - (w_A + K)A - w_B B + \psi(t,A) - \phi(t,B) \\ \textit{subject to constraint } B \geq 0\end{gathered} \tag{9}$$

Claim 3. Under the same conditions as in Claim 1, a tax of K is levied on every A-worker. Increasing K would have the effect of:

1. Increasing the wage and employment of B-workers and decreasing the wage and employment of A-workers.
2. Shrinking the set of employers who avoid B-workers.
3. Depressing total employment.

The proof is in the appendix.

The intuition and discussion regarding Claim 3 is already contained in that following Claim 1. If taxing B-workers, by the antidiscrimination fine,

hurts them and benefits the A-workers, the reverse is what one needs. Taxing the A-workers, instead, would hurt them and benefit the B-workers. When one type of workers is taxed, as in both cases, total employment suffers and government revenue increases.

5. EFFECTIVE ALTERNATIVE POLICIES: SUBSIDY

Now examine a policy of subsidizing the disfavored workers. It also is easy to enforce but is costly and burdens the government's budget.

Suppose a subsidy H is paid for each B-worker employed. Now employer t's target is

$$\underset{A,B}{\text{Max}}\, U = f(A+B) - w_A A - (w_B - H)B + \psi(t,A) - \phi(t,B) \quad \textit{subject to constraint } B \geq 0 \tag{10}$$

Claim 4. Under the same conditions as in Claim 1, a subsidy of H is paid for every B-worker. Increasing H would have the effect of:

1. Increasing the wage and employment of B-workers and decreasing the wage and employment of A-workers.
2. Shrinking the set of employers who avoid B-workers.
3. Boosting total employment.

The proof is in the appendix.

Intuition and discussion:
A subsidy is a negative tax, so its effect is as that of a tax, only in the reverse direction. Consequently, this law/policy is an identical mirror image of that discussed in Claim 1.

This policy seems an unmitigated blessing, the perfect solution: employment and wages of the disfavored workers increase, employment and wages of the favored workers decrease, total employment increases. On the other hand, the weakness of this policy is that it requires public financing and is a drain on the budget.

6. QUOTAS

Instead of attempting to curb discrimination by fines, a legislature could try quotas as a means of correcting the wrong. One could argue that affirmative

action resembles better a quota policy than fines. Therefore, it is pertinent to analyze the effectiveness of such policy.

The situation is somewhat similar to international trade where a protective policy differentiates between foreign goods and their domestic substitutes. There also one can use import quotas or differential taxation, i.e., import duties or subsidy for domestic production. It is well known that for any result achieved by a quota, one can design a tax that would attain it as well. The difference is in the administrative cost of the two programs and distributional aspects. Whereas the tax revenue goes to the government, a quota produces a rent endowed by its holders.

Fig. 2 demonstrates the consequences of a quota of A-workers in their labor market. Suppose only A_0 such workers are allowed. Competition among employers for the limited number of workers would raise their wage to OF. Competition among the workers vying for employment would reduce the wage to OE. The difference EF per worker is collected by those privileged to own a permit, be it the A-worker himself, an employer or a go between. A tax of EF would attain the same equilibrium in the A- and B-labor markets with the rent going to the government.

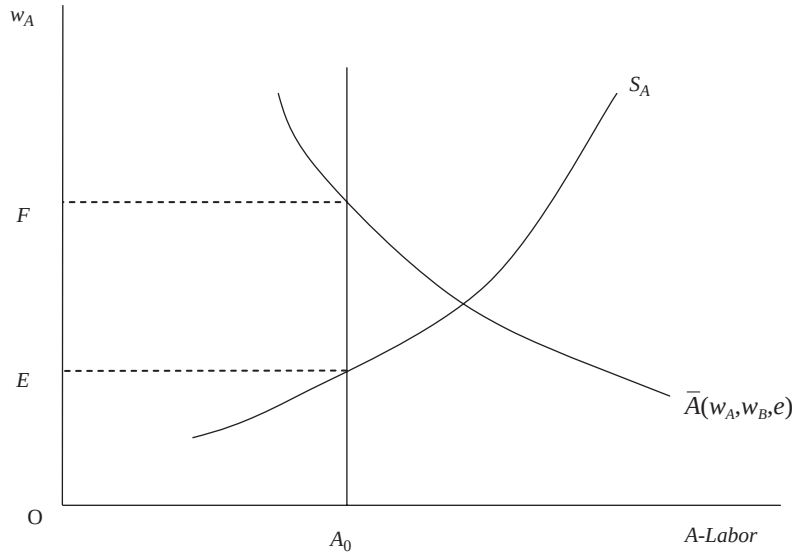

Fig. 2. Demand, Supply and Quota in the A-Labor Market.

7. CONCLUDING REMARKS

Using Becker's 'taste for discrimination' model, the chapter analyzes the current legislation against wage discrimination and finds it counter-productive. Using a costly apparatus of auditing, detecting and fining violators does not deliver the desired results. If the fine is reimbursed to the disadvantaged workers in order to undo the discrimination, it equally affects the demand for and the supply of *B*-workers, because their expected wage includes the fine, and has no real effect. If the fine is collected and kept, it shifts employment away from the workers it seeks to help, depressing their wage.

In contrast, two alternative policies are proposed that would curb discrimination in the labor market by shifting the equilibrium wage levels toward equality. One is a tax on the favored workers and the other is a subsidy for disfavored workers. Varying the tax rate K, or the subsidy rate H, the wage differential may be lowered, abolished and even reversed. However, the tax, like any tax, depresses total employment while the subsidy does the opposite. The two are also opposites with regard to costs.

A quota of favored workers may serve as a substitute for a tax on them but with possibly undesirable distributional effect and costly administration. Affirmative action is, in essence, some hybrid of a tax on employing *A*-type workers and a quota, only administered in an indirect, clumsy and costly way. Still, that explains its humble impact in the right direction (see, e.g., Leonard, 1996). An explicit and direct tax would do much more and with negative costs.

NOTE

1. The term affirmative action describes policies aimed at weak or discriminated groups concerning employment or education.

ACKNOWLEDGMENTS

We thank Gideon Yaniv for his helpful remarks. We wish to thank two anonymous referees for helpful comments and suggestions. Remaining errors are ours.

REFERENCES

Abram, T. G. (1993). The law, its interpretation, levels of enforcement activity and effect on employer behavior. *American Economic Review*, *83*, 62–66.

Acemoglu, D., & Angrist, J. (2001). Consequences of employment protection: The case of the Americans with disabilities act. *Journal of Political Economy*, *109*, 915–957.

Arrow, K. (1972a). Models of job discrimination. In: A. H. Pascal (Ed.), *Racial discrimination in economic life* (pp. 83–102). Lexington, MA: Lexington Books.

Arrow, K. (1972b). Some mathematical models of race discrimination in the labor market. In: A. H. Pascal (Ed.), *Racial discrimination in economic life* (pp. 187–204). Lexington, MA: Lexington Books.

Arrow, K. J. (1974). The theory of discrimination. In: O. Ashenfelter & A. Rees (Eds), *Discrimination in the labor markets*. Princeton, NJ: Princeton University Press.

Becker, G. S. (1971). *The economics of discrimination* (Original edition, 1957). Chicago, IL: University of Chicago Press.

Chang, Y. M., & Ehrlich, I. (1985). On the economics of compliance with the minimum wage law. *Journal of Political Economy*, *93*, 84–91.

Donohue, J. J., & Siegelman, P. (1991). The changing nature of employment discrimination litigation. *Stanford Law Review*, *43*, 983–1033.

Heckman, J. J., & Payner, B. S. (1989). Determining the impact of federal antidiscrimination policy on the economic status of Blacks. *American Economic Review*, *79*, 138–177.

Leonard, J. S. (1996). Wage disparities and affirmative action in the 1980s. *American Economic Review*, *86*, 285–289.

Oyer, P., & Schaefer, S. (2000). Layoffs and litigation. *RAND Journal of Economics*, *31*, 345–358.

Oyer, P., & Schaefer, S. (2002). Sorting quotas, and the civil rights act of 1991: Who hires when it's hard to fire?. *Journal of Law and Economics*, *45*, 41–68.

Posner, R. A. (1987). The efficiency and efficacy of title VII. *University of Pennsylvania Law Review*, *136*, 513–519.

Tobol, Y. (2005). Wage discrimination as an illegal behavior. *Economics Bulletin*, *10*(1), 1–10.

Yaniv, G. (2001). Minimum wage noncompliance and the employment decision. *Journal of Labor Economics*, *19*, 596–603.

APPENDIX

Proof of Claim 1. Totally differentiating Eqs. (5) and (6) with respect to e, one gets an interior solution

$$f''(A+B)\left(\frac{\mathrm{d}A}{\mathrm{d}e}+\frac{\mathrm{d}B}{\mathrm{d}e}\right)-\frac{\mathrm{d}w_A}{\mathrm{d}e}+\psi''(t,A)\frac{\mathrm{d}A}{\mathrm{d}e}=0 \tag{A.1}$$

$$\begin{aligned} f''(A+B)\left(\frac{\mathrm{d}A}{\mathrm{d}e}+\frac{\mathrm{d}B}{\mathrm{d}e}\right)-\frac{\mathrm{d}w_B}{\mathrm{d}e}-\phi''(t,B)\frac{\mathrm{d}B}{\mathrm{d}e} \\ -(w_A-w_B)-e\left(\frac{\mathrm{d}w_A}{\mathrm{d}e}-\frac{\mathrm{d}w_B}{\mathrm{d}e}\right)=0 \end{aligned} \tag{A.2}$$

From Eq. (A.1) one gets

$$\left(f''(A+B)+\psi''(t,A)\right)\frac{\mathrm{d}A}{\mathrm{d}e}-\frac{\mathrm{d}w_A}{\mathrm{d}e}=-f''(A+B)\frac{\mathrm{d}B}{\mathrm{d}e} \tag{A.3}$$

Some employers use both types (interior solution) and some only A types (corner solution). We here assume, and later show, that the signs of $\mathrm{d}\bar{A}/\mathrm{d}e$ and $\mathrm{d}\bar{B}/\mathrm{d}e$ are equal to $\mathrm{d}A/\mathrm{d}e$ and $\mathrm{d}B/\mathrm{d}e$ of the employers in the first set, respectively. Now, because the labor-supply curves in Eq. (7) are rising in wage, i.e., $(\mathrm{d}S_A/\mathrm{d}w_A)>0$ and $(\mathrm{d}S_B/\mathrm{d}w_B)>0$, changes in wage and in total quantity for A-workers must be of the same sign as can be seen by differentiating Eq. (7):

$$\frac{\mathrm{d}S_A}{\mathrm{d}w_A}\frac{\mathrm{d}w_A}{\mathrm{d}e}=\frac{\mathrm{d}\bar{A}(w_A,w_B,e)}{\mathrm{d}e}=\int_0^1\frac{\mathrm{d}A(t,w_A,w_B,e)}{\mathrm{d}e}g(t)\mathrm{d}t \tag{A.4}$$

and similarly for B:

$$\frac{\mathrm{d}S_B}{\mathrm{d}w_B}\frac{\mathrm{d}w_B}{\mathrm{d}e}=\frac{\mathrm{d}\bar{B}(w_A,w_B,e)}{\mathrm{d}e}=\int_0^1\frac{\mathrm{d}B(t,w_A,w_B,e)}{\mathrm{d}e}g(t)\mathrm{d}t \tag{A.5}$$

Therefore, Eq. (A.3) implies that $\mathrm{d}A/\mathrm{d}e$ and $\mathrm{d}B/\mathrm{d}e$ have opposite signs. Now, $\mathrm{d}B/\mathrm{d}e$ must be negative, otherwise one gets into a contradiction. Substituting Eqs. (A.1) in (A.2), one gets

$$\frac{\mathrm{d}w_A}{\mathrm{d}e}-\frac{\mathrm{d}w_B}{\mathrm{d}e}=\frac{(w_A-w_B)+\psi''(t,A)(\mathrm{d}A/\mathrm{d}e)+\phi''(t,A)(\mathrm{d}B/\mathrm{d}e)}{1-e} \tag{A.6}$$

If $(\mathrm{d}B/\mathrm{d}e)>0$ the left-hand side of Eq. (A.6) is negative because then from Eq. (A.5) $(\mathrm{d}w_B/\mathrm{d}e)>0$ and by the opposite sign $(\mathrm{d}A/\mathrm{d}e)<0$, $(\mathrm{d}w_A/\mathrm{d}e)<0$ while the right-hand side is positive. As for the employers with $B=0$, Eq. (A.2) is irrelevant, $(\mathrm{d}B/\mathrm{d}e)=0$ and Eq. (A.3) implies that $\mathrm{d}A/\mathrm{d}e$ is opposite in sign to $\mathrm{d}w_A/\mathrm{d}e$. Still, the first set of employers has the upper hand in their impact on the employment of A-workers. Otherwise, if $(\mathrm{d}A/\mathrm{d}e)<0$, it would contradict Eq. (A.3) where the signs of $\mathrm{d}B/\mathrm{d}e$, $\mathrm{d}A/\mathrm{d}e$ must be opposite. So one may conclude that $(\mathrm{d}\bar{B}/\mathrm{d}e)<0$, $(\mathrm{d}\bar{A}/\mathrm{d}e)>0$, $(\mathrm{d}w_B/\mathrm{d}e)<0$, $(\mathrm{d}w_A/\mathrm{d}e)>0$. The decline of total employment may be deduced from Eq. (A.1):

$$\frac{\mathrm{d}A}{\mathrm{d}e}+\frac{\mathrm{d}B}{\mathrm{d}e}=\frac{(\mathrm{d}w_A/\mathrm{d}e)-\psi''(t,A)(\mathrm{d}A/\mathrm{d}e)}{f''(A+B)}<0.$$

For the critical employer t^* the equality

$$f'(A(t^*))-w_B-e(w_A-w_B)-\phi'(t^*,0)=0 \tag{A.7}$$

holds, so that for $t^* < t$ all employers choose $B = 0$. Differentiating Eq. (A.7), including t^*, with respect to e, one gets

$$\frac{\mathrm{d}t^*}{\mathrm{d}e} = \frac{f''(A(t^*))(\mathrm{d}A/\mathrm{d}e) - (w_A - w_B) - e((\mathrm{d}w_A/\mathrm{d}e) - (\mathrm{d}w_B/\mathrm{d}e)) - (\mathrm{d}w_B/\mathrm{d}e)}{(\mathrm{d}\phi'(t,0)/\mathrm{d}t)} < 0 \tag{A.8}$$

because the numerator is negative from Eq. (A.1) and the denominator is positive from the assumption of ascending order of preferences. Since t^* is reduced, the interval of employers who shun B-workers, $[t^*,1]$, expands.

Proof of Claim 2. One can write Eq. (6) as

$$\frac{\partial(EU)}{\partial B} = f'(A+B) - [ew_A + (1-e)w_B] - \phi'(t,B) = 0 \tag{A.9}$$

That is, the demand of each firm depends on the expected wage exactly like the supply of B. So the B-labor market equilibrium, the equivalent of Eq. (7), can now be written as

$$S_B(ew_A + (1-e)w_B) = \int_0^1 B(t, ew_A + (1-e)w_B, e)g(t)\mathrm{d}t \tag{A.10}$$

A stronger enforcement equally shifts down the demand and the supply for B-workers in the (B, w_B) plane. Suppose e goes up from e_1 to e_2. The market equilibrium will persist at the same quantities A, B and w_A if the B-wage is adjusted down so as to hold the same expectation. That is,

$$e_2 w_A + (1-e_2)\bar{w}_B = e_1 w_A + (1-e_1)w_B \quad \text{or} \quad \bar{w}_B = \frac{(1-e_1)}{(1-e_2)}w_B - \frac{(e_2-e_1)}{(1-e_2)}w_A \tag{A.11}$$

Formally, the equivalent of Eq. (A.5) now has the form:

$$\begin{aligned} &\frac{\mathrm{d}S_B}{\mathrm{d}[ew_A + (1-e)w_B]} \frac{\mathrm{d}[ew_A + (1-e)w_B]}{\mathrm{d}e} \\ &\quad = \frac{\mathrm{d}S_B}{\mathrm{d}[ew_A + (1-e)w_B]} \left[w_A - w_B + e\left(\frac{\mathrm{d}w_A}{\mathrm{d}e} - \frac{\mathrm{d}w_B}{\mathrm{d}e}\right) + \frac{\mathrm{d}w_B}{\mathrm{d}e} \right] \\ &\quad = \int_0^1 \frac{\mathrm{d}B(t, ew_A + (1-e)w_B)}{\mathrm{d}e} g(t)\mathrm{d}t \end{aligned} \tag{A.12}$$

From the above considerations, the solution to the system of Eqs. (A.1), (A.2), (A.4) and (A.12) is

$$\frac{\mathrm{d}A}{\mathrm{d}e}=0,\ \frac{\mathrm{d}B}{\mathrm{d}e}=0,\ \frac{\mathrm{d}w_A}{\mathrm{d}e}=0,\ \frac{\mathrm{d}w_B}{\mathrm{d}e}=-\frac{w_A-w_B}{1-e} \tag{A.13}$$

Proof of Claim 3. Similar to the analysis in Section 2, the first-order conditions for a maximum are

$$\frac{\partial U}{\partial A}=f'(A+B)-w_A-K+\psi'(t,A)=0 \tag{A.14}$$

$$\frac{\partial U}{\partial B}=f'(A+B)-w_B-\phi'(t,B)=0 \tag{A.15}$$

Totally differentiating Eqs. (A.14) and (A.15) with respect to K

$$f''(A+B)\left(\frac{\mathrm{d}A}{\mathrm{d}K}+\frac{\mathrm{d}B}{\mathrm{d}K}\right)-\frac{\mathrm{d}w_A}{\mathrm{d}K}-1+\psi''(t,A)\frac{\mathrm{d}A}{\mathrm{d}K}=0 \tag{A.16}$$

$$f''(A+B)\left(\frac{\mathrm{d}A}{\mathrm{d}K}+\frac{\mathrm{d}B}{\mathrm{d}K}\right)-\frac{\mathrm{d}w_B}{\mathrm{d}K}-\phi''(t,B)\frac{\mathrm{d}B}{\mathrm{d}K}=0 \tag{A.17}$$

From Eqs. (A.16) and (A.17) one derives an interior solution

$$\frac{\mathrm{d}w_A}{\mathrm{d}K}-\frac{\mathrm{d}w_B}{\mathrm{d}K}=-1+\psi''(t,A)\frac{\mathrm{d}A}{\mathrm{d}K}+\phi''(t,A)\frac{\mathrm{d}B}{\mathrm{d}K} \tag{A.18}$$

and also

$$(f''(A+B)+\psi''(t,A))\frac{\mathrm{d}A}{\mathrm{d}K}-\frac{\mathrm{d}w_A}{\mathrm{d}K}=1-f''(A+B)\frac{\mathrm{d}B}{\mathrm{d}K} \tag{A.19}$$

$$f''(A+B)\frac{\mathrm{d}A}{\mathrm{d}K}=\frac{\mathrm{d}w_B}{\mathrm{d}K}-(f''(A+B)-\phi''(t,A))\frac{\mathrm{d}B}{\mathrm{d}K} \tag{A.20}$$

which imply, recalling that the supply curves are positively sloped so $(\mathrm{d}A/\mathrm{d}K)(\mathrm{d}w_A/\mathrm{d}K)>0$, $(\mathrm{d}B/\mathrm{d}K)(\mathrm{d}w_B/\mathrm{d}K)>0$, that $\mathrm{d}A/\mathrm{d}K$ and $\mathrm{d}B/\mathrm{d}K$ have opposite signs while Eq. (A.18) shows that the signs must be $(\mathrm{d}A/\mathrm{d}K)<0, (\mathrm{d}w_A/\mathrm{d}K)<0, (\mathrm{d}B/\mathrm{d}K)>0, (\mathrm{d}w_B/\mathrm{d}K)>0$. In this case the corner-solution employers react the same way: Eq. (A.17) is irrelevant for them, $(\mathrm{d}B/\mathrm{d}K)=0$ and $(\mathrm{d}A/\mathrm{d}K)<0, (\mathrm{d}w_A/\mathrm{d}K)<0$ from Eq. (A.16).

From Eq. (A.17) one concludes that

$$\frac{\mathrm{d}A}{\mathrm{d}K}+\frac{\mathrm{d}B}{\mathrm{d}K}=\frac{(\mathrm{d}w_B/\mathrm{d}K)+\phi''(t,B)(\mathrm{d}B/\mathrm{d}K)}{f''(A+B)}<0 \tag{A.21}$$

Finally, for the critical employer t^*

$$f'(A(t^*))-w_B-\phi'(t^*,0)=0 \tag{A.22}$$

so that for $t^*<t$ all employers choose $B=0$. Differentiating Eq. (A.22), including t^*, with respect to K, one gets

$$\frac{\mathrm{d}t^*}{\mathrm{d}K}=\frac{f''(A(t^*))(\mathrm{d}A/\mathrm{d}K)-(\mathrm{d}w_B/\mathrm{d}K)}{(d\phi'(t,0))/\mathrm{d}t}>0 \tag{A.23}$$

The numerator of Eq. (A.23) is positive from Eq. (A.20) and the denominator is positive from the assumption of ascending order of preferences. Since t^* increases, the interval of employers who shun B-workers, $[t^*,1]$, shrinks.

Proof of Claim 4. Similar to the analysis in Section 4, the first-order conditions for a maximum are

$$\frac{\partial U}{\partial A}=f'(A+B)-w_A+\psi'(t,A)=0 \tag{A.24}$$

$$\frac{\partial U}{\partial B}=f'(A+B)-w_B+H-\phi'(t,B)=0 \tag{A.25}$$

Totally differentiating Eqs. (A.24) and (A.25) with respect to H, yields:

$$f''(A+B)\left(\frac{\mathrm{d}A}{\mathrm{d}H}+\frac{\mathrm{d}B}{\mathrm{d}H}\right)-\frac{\mathrm{d}w_A}{\mathrm{d}H}+\psi''(t,A)\frac{\mathrm{d}A}{\mathrm{d}H}=0 \tag{A.26}$$

$$f''(A+B)\left(\frac{\mathrm{d}A}{\mathrm{d}H}+\frac{\mathrm{d}B}{\mathrm{d}H}\right)-\frac{\mathrm{d}w_B}{\mathrm{d}H}+1-\phi''(t,B)\frac{\mathrm{d}B}{\mathrm{d}H}=0 \tag{A.27}$$

From these equations one derives

$$\frac{\mathrm{d}w_A}{\mathrm{d}H}-\frac{\mathrm{d}w_B}{\mathrm{d}H}=-1+\psi''(t,A)\frac{\mathrm{d}A}{\mathrm{d}H}+\phi''(t,A)\frac{\mathrm{d}B}{\mathrm{d}H} \tag{A.28}$$

From Eqs. (A.26) and (A.27) one gets

$$(f''(A+B)+\psi''(t,A))\frac{\mathrm{d}A}{\mathrm{d}H}-\frac{\mathrm{d}w_A}{\mathrm{d}H}=-f''(A+B)\frac{\mathrm{d}B}{\mathrm{d}H} \tag{A.29}$$

$$f''(A+B)\frac{\mathrm{d}A}{\mathrm{d}H}=-1+\frac{\mathrm{d}w_B}{\mathrm{d}H}-(f''(A+B)-\phi''(t,A))\frac{\mathrm{d}B}{\mathrm{d}H} \tag{A.30}$$

which imply that $\partial A/\partial H$ and $\partial B/\partial H$ have opposite signs, while Eq. (A.28) shows that the signs must be $(\mathrm{d}A/\mathrm{d}H)<0, (\mathrm{d}w_A/\mathrm{d}H)<0$, $(\mathrm{d}B/\mathrm{d}H)>0, (\mathrm{d}w_B/\mathrm{d}H)>0$, and $(\mathrm{d}A/\mathrm{d}K)<0, (\mathrm{d}w_A/\mathrm{d}K)<0, (\mathrm{d}B/\mathrm{d}K)>0$, $(\mathrm{d}w_B/\mathrm{d}K)>0$.

In this case also the corner-solution employers react the same way as others: Eq. (A.27) is irrelevant for them, $(\mathrm{d}B/\mathrm{d}H)=0$ and from Eq. (A.26) $(\mathrm{d}A/\mathrm{d}H)<0, (\mathrm{d}w_A/\mathrm{d}H)<0$. From Eq. (A.26) one concludes that

$$\frac{\mathrm{d}A}{\mathrm{d}H}+\frac{\mathrm{d}B}{\mathrm{d}H}=\frac{(\mathrm{d}w_A/\mathrm{d}H)-\psi''(t,B)(\mathrm{d}A/\mathrm{d}H)}{f''(A+B)}>0 \tag{A.31}$$

Finally, for the critical employer t^* from Eq. (A.27), differentiating it, including t^*, with respect to H, one gets

$$\frac{\mathrm{d}t^*}{\mathrm{d}H}=\frac{f''(A(t^*))(\mathrm{d}A/\mathrm{d}H)-(\mathrm{d}w_B/\mathrm{d}H)}{(d\phi'(t,0)/\mathrm{d}t)}>0 \tag{A.32}$$

The numerator of Eq. (A.31) is positive from Eq. (A.30) and the denominator is positive from the assumption of ascending order of preferences. Since t^* increases, the interval of employers who shun B-workers, $[t^*,1]$, shrinks.

PATTERNS OF NOMINAL AND REAL WAGE RIGIDITY ☆

Louis N. Christofides and Paris Nearchou

ABSTRACT

We study the distortions that downward nominal and real wage rigidity would induce to a flexible form of a notional, rigidity-free, distribution of wage change using the histogram-location approach. We examine alternative methods of generating the histograms that support the econometric search for rigidity distortions and implement our approach to inflation sub-periods that should be characterised by different patterns of nominal and real rigidities. We establish the general applicability of the approach to these sub-periods and find results consistent with expectations.

1. INTRODUCTION

In Keynesian models, the notion of downward nominal wage rigidity (DNWR) plays an important role in 'rationalising' the failure of the labour

☆Earlier versions of this work were presented by Christofides at the Banque de France conference on Wage Bargaining, Employment, and Monetary and Economic Policies, October 9–10, 2007, in Paris, and the conference in honour of Ray Rees, July 3–4, 2008, in Munich, and by Nearchou at the September 2008 EALE conference in Amsterdam.

Jobs, Training and Worker Well-Being
Research in Labor Economics, Volume 30, 301–337
Copyright © 2010 by Emerald Group Publishing Limited
All rights of reproduction in any form reserved
ISSN: 0147-9121/doi:10.1108/S0147-9121(2010)0000030013

market to clear and the existence of unemployment. In these models, employment is determined along the labour demand curve (the 'short' side of the market) and, since increases in employment require movements along this demand curve, the real wage behaves counter-cyclically. Early attempt by Dunlop (1938) and Tarshis (1939), but also more recent papers by Solon, Barsky, and Parker (1994) and Abraham and Haltiwanger (1995) examine nominal wage rigidity indirectly by looking at the cyclical properties of the real wage rate.[1]

McLaughlin's (1994) paper shifted attention to wage growth distributions (WGDs) for individuals in panel data, thus giving rise to a more inductive approach to this issue: What do data on the earnings of individuals over time imply about wage rigidity? The papers by, inter alia, Lebow, Stockton, and Wascher (1995); Fortin (1996); Kahn (1997); Card and Hyslop (1997); Crawford and Harrison (1998); Smith (2000); Altonji and Devereux (2000); Christofides and Stengos (2001, 2002, 2003); Christofides and Leung (2003); Christofides and Li (2005); Dickens and Groshen (2004) and Holden and Wulfsberg (2007) all follow this broad approach. Some of these papers deal not only with DNWR, but also with aspects of real wage rigidity.

The extent to which DNWR and downward real wage rigidity (DRWR) co-exist and interact are points that are worth investigating further.[2] Papers which address both types of rigidity based on the maximum-likelihood approach (e.g. Bauer, Bonin, Goette, & Sunde, 2007; Barwell & Schweitzer, 2007) assume that some agents are subject to DNWR, some to DRWR and some to neither type of downward wage rigidity (DWR), specifying the effects of each separately and leaving the overall picture to be determined by a mixing process. In Christofides and Nearchou (2007), we describe how the 'histogram-location' approach, that goes back to Kahn (1997), can be modified to detect not only DNWR but also DRWR. Our approach makes no parametric assumptions and does not allocate individuals to rigidity regimes, as is inherent in the likelihood-based literature. It does rely, for the identification of possible DRWR effects, on having an inflation experience which is sufficiently diverse to allow the median of the WGD to differ from the centre of the anticipated inflation distribution (AID). As the two points of central tendency drift apart, distortions to the WGD around the mean of the AID (the expected inflation rate or $\dot{P}^e$) may be detected and, if consistent with a priory restrictions, these distortions can be attributed to DRWR. DNWR is still investigated as in Kahn (1997) and Christofides and Leung (2003) by focusing on distortions in the WGD at the point zero. Thus both types of DWR can be examined. Christofides and Nearchou (2007) implement this model using wage contract data from Canada over a period

(1976–1999) which is characterised by very high inflation, moderate inflation as well as extremely low inflation; however, it is applicable to any data set with a panel dimension and to any inflation period, provided sufficient care is taken in specifying the model.

In this chapter, we extend this earlier work in several directions. First, the histograms are defined such that it is the median of the WGD that is located in the middle of the bin that contains it, rather than the point zero in its respective bin, as in our earlier paper. In that paper, the focus on zero allowed a neater exploration of the possibility of menu costs. Since this particular type of rigidity mechanism, while statistically significant, appears to account for approximately one percentage point of distortion in the WGD, we now wish to explore a possible lack of clarity that may arise when the median of the WGD is not centred in its so-called 'median' bin for each and every year in the sample. A second issue that we now address is the extent to which the relative frequency approach, used in our earlier work to construct the 'stage 1' bin heights that underlie our 'stage 2' econometric exploration for DWR effects, can be improved by using non-parametric kernel methods. This should be the case on a priori grounds, given that the relative frequency approach essentially imposes a zero bandwidth, rather than choosing one optimally. Finally, having dealt with these 'stage 1' issues, we explore the existence of and possible interactions between DNWR and DRWR by paying special attention to inflation sub-periods where (i) one kind of rigidity may be present while the other may not, as in periods of high inflation where DNWR may be not be relevant (ii) both types of rigidity may be important but DRWR may be more important than DNWR, as in periods of moderate inflation and (iii) both types or rigidity may be important but DNWR may be more important than DRWR, as during the more recent and prolonged period of extremely low inflation. These explorations raise technical concerns about how the model might be implemented during various sub-periods and they shed light on how these rigidities operate. They also suggest how our approach might be tailored to samples from countries and periods that share generic features with the sub-samples examined here.

Section 2 considers how the notions of DWR are implemented in our work; it also considers each of the three points raised in the previous paragraph, thus better-motivating the contribution of the present chapter. Section 3 examines relevant features of the contract data that are used both here and in earlier work; working with the same database allows useful comparisons. Section 4 presents the econometric specification used and its application to inflation sub-periods. Section 5 presents the results obtained

and Section 6 summarises our findings and explores possible further work that might be undertaken.

2. MOTIVATING ISSUES

In all our work thus far, we assume that, *if* DNWR holds, agents would be reluctant to accept a nominal wage cut and instead would settle for a nominal wage freeze. At the population level, this reluctance would mean fewer cuts in nominal wages and more nominal wage freezes *relative to* the case of no rigidity. In terms of the distribution of nominal wage growth rates, this translates into a shift of probability mass from negative values of the support of the WGD to the point zero. Therefore, the rigidity-contaminated nominal WGD would show a deficit of probability mass for negative values of the support, and a surplus at the point zero, relative to the notional distribution. At the same time, the two distributions should be identical beyond the point zero. Justifications for nominal rigidity range from the comparability and fairness arguments documented in Bewley (1999) to the theoretical papers by Macleod and Malcomson (1993), Malcomson (1997) and Holden (1994, 2004) which build on the notion that nominal wages can be changed only by mutual consent.

To the extent that agents perceive that small price changes (positive or negative) are not worth the cost of implementing them, some deficit, which may not be symmetric, may appear in the area of the actual WGD immediately below and above zero. In this paper, we still check for these effects when the whole period is considered.

DRWR can be defined in a similar way to DNWR. We assume that DRWR describes the situation where agents are reluctant to accept real-wage cuts but instead would settle for a real wage freeze. In practice, this attitude takes the form of reluctance towards accepting reductions in the *anticipated* real wage since, at the time of bargaining, future inflation is unknown. As in the case of DNWR, the presence of DRWR would distort the shape of the nominal WGD. At the population level, this would mean that agents who face nominal wage growth at a rate below anticipated inflation would settle for a nominal wage increase equal to the anticipated rate of inflation. Consequently, the presence of DRWR would shift probability mass to the right, from smaller values of nominal wage growth towards the values of anticipated inflation in the population.

The exact form of the shift of mass to the right towards the values of anticipated inflation depends on the nature of the rigidity mechanism and

the joint distribution of the notional (nominal) wage growth and anticipated inflation among all agents. Nevertheless, without any distributional assumptions, it is possible to distinguish three regions in the nominal WGD for which we can make qualitative predictions about the nature of the distortions. For simplicity, suppose that the support for the AID lies inside that for the WGD. First, the interval of values that lies to the left of the support of the AID could only loose mass to the right, since all agents whose nominal wage growth falls in this region face the prospect of a real wage cut. Therefore, in this region, the rigidity-contaminated distribution can only exhibit a deficit. Second, the interval of values that lies to the right of the support of the distribution of anticipated inflation would not be distorted, since all agents whose nominal wage growth falls in this region face the prospect of a real wage increase. Third, the interval of values that corresponds to the support of the AID, will attract mass from its left, and therefore for this interval the rigidity-contaminated distribution will exhibit a surplus in total. However, it is possible that, in some parts of this interval, the rigidity-contaminated distribution will exhibit a deficit. In terms of the probability histogram, this is because a particular bin that coincides with values of anticipated inflation can attract mass from bins to its left but at the same time loose mass to bins to its right that also coincide with values of anticipated inflation. The net effect cannot be clear without knowledge of how notional wage growth and anticipated inflation are jointly distributed. The only exception is the rightmost bin in this region, for which we know that it cannot exhibit a deficit since all other bins that contain values of anticipated inflation lie to its left. Despite this uncertainty, we could assume that it would be more likely that bins that lie further to the left in this interval will show a deficit and bins further to the right will show a surplus. The sum of the net effects to the maximum point of the AID support should be zero.

This discussion indicates that the search for DRWR effects is inherently much more difficult than that for DNWR. The distortions arising from DRWR are potentially spread over a wide range of the WGD, beginning with the minimum point of the support and up to the maximum point of the AID. A further complication is that the precise limits of the support of the AID can only be conjectured; it is possible that it extends well to the left and right of $\dot{P}^e$ so that the transfer of mass may involve several bins on either side of $\dot{P}^e$. It is more likely, however, that more bins to the left will be involved than bins to the right given our discussion above.

It is also interesting to note what the presence of DRWR means for the distribution of *actual* real-wage growth. If we accept that typically the AID extends below and above the realised inflation value, then the presence of

DRWR is consistent with observing real-wage cuts (relative to the realised value of inflation), even in the case of absolute (i.e. complete reluctance by all to accept a real-wage cut) DRWR. Therefore, the occurrence of real wage cuts does not, in general, suggest that DRWR does not exist; real-wage cuts are inconsistent only with the case of absolute DRWR and perfect foresight.

Having outlined our broad approach to how we expect DWR to impact on the WGD, we now turn to the three main issues that we wish to explore in this chapter.

2.1. Centering on the Median

In our earlier work, the wage change information was used, in stage 1, to construct histograms with bins located such that the point zero was at the centre of the bin that contained it, so as to facilitate the exploration of possible menu-cost behaviour. Suppose that, in this zero-based construction, the median of the WGD was only just large enough to enter the so-called 'median' bin. Then the bin containing the point zero (at its centre) might be located at the $-j$th bin, that is j bins below the 'median' bin. If, on the other hand, the histograms for each year are constructed with the 'median' bin centered on the actual median for the year, then in the above example the bin containing the point zero (not at its centre) will still be the $-j$th one. However, any other arbitrary point in the WGD support *could* belong to one bin under the zero-based construction and to an adjacent bin under median centering. An important such point is $\widehat{\dot{P}^e}$, since it figures prominently in our search for DRWR distortions. While, in practice, we do not expect these difficulties to be severe, it is preferable to now construct the yearly histograms by centering on the yearly median. The histograms presented in Figs. 1–3 below are indeed constructed in this manner and are very similar with the zero-based ones presented, for selected years, in Christofides and Nearchou (2007). Note that the bins containing the expected inflation rate and the point zero are indicated in the three figures.

2.2. Kernel Estimates of Histogram Heights

Our test procedures involve comparisons between the notional (DWR-free) and the actual (rigidity-contaminated) WGD. These comparisons are carried out using probability histograms. We divide the support of the actual WGD into sub-intervals (bins) and compare the amount of

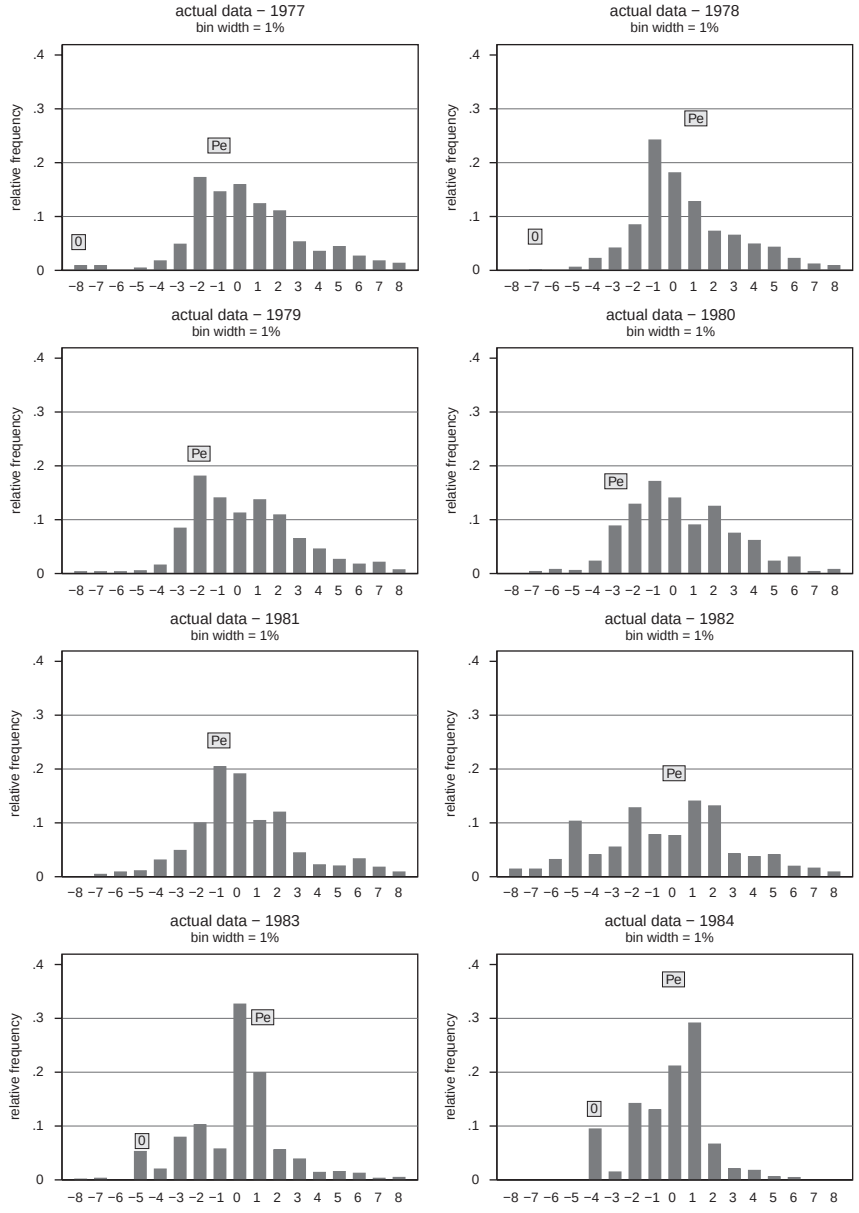

Fig. 1. Standardised (Median-Centred) Relative Frequency Histograms.

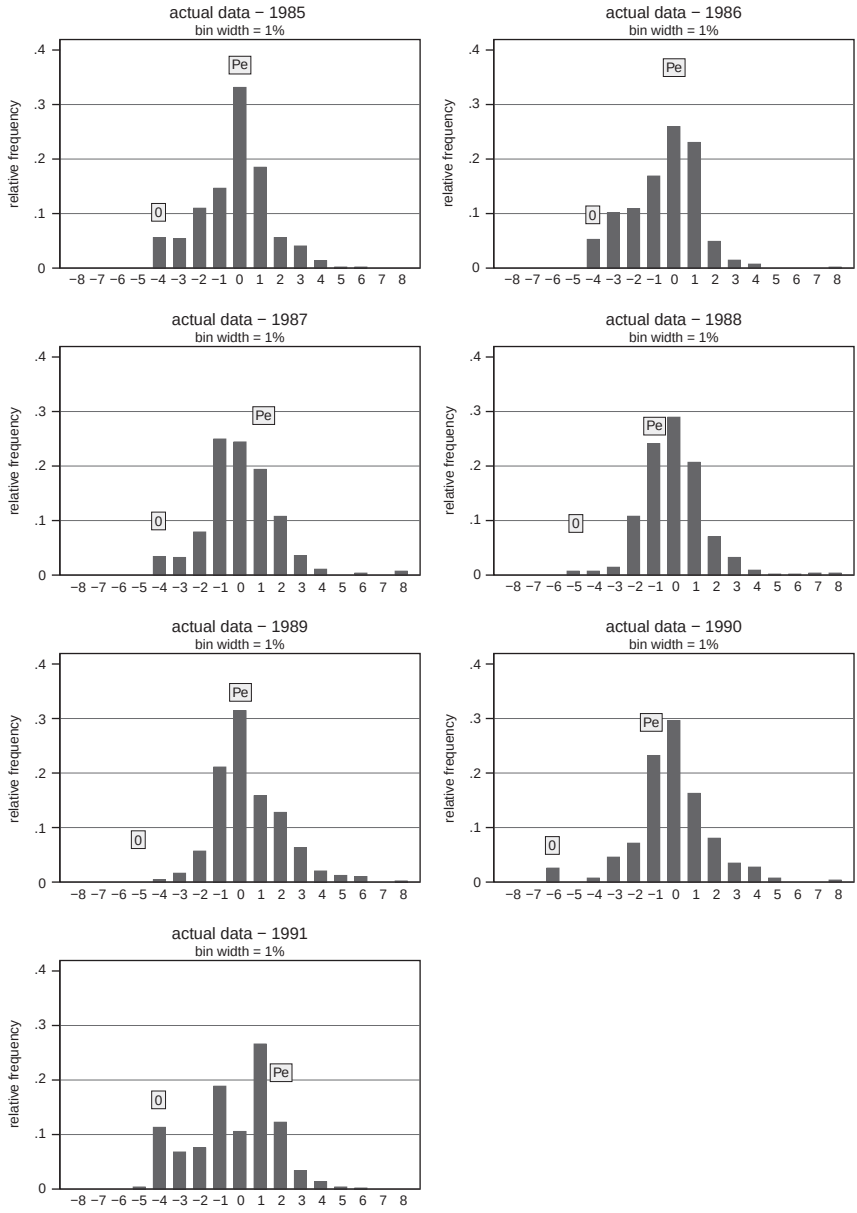

Fig. 2. Standardised (Median-Centred) Relative Frequency Histograms.

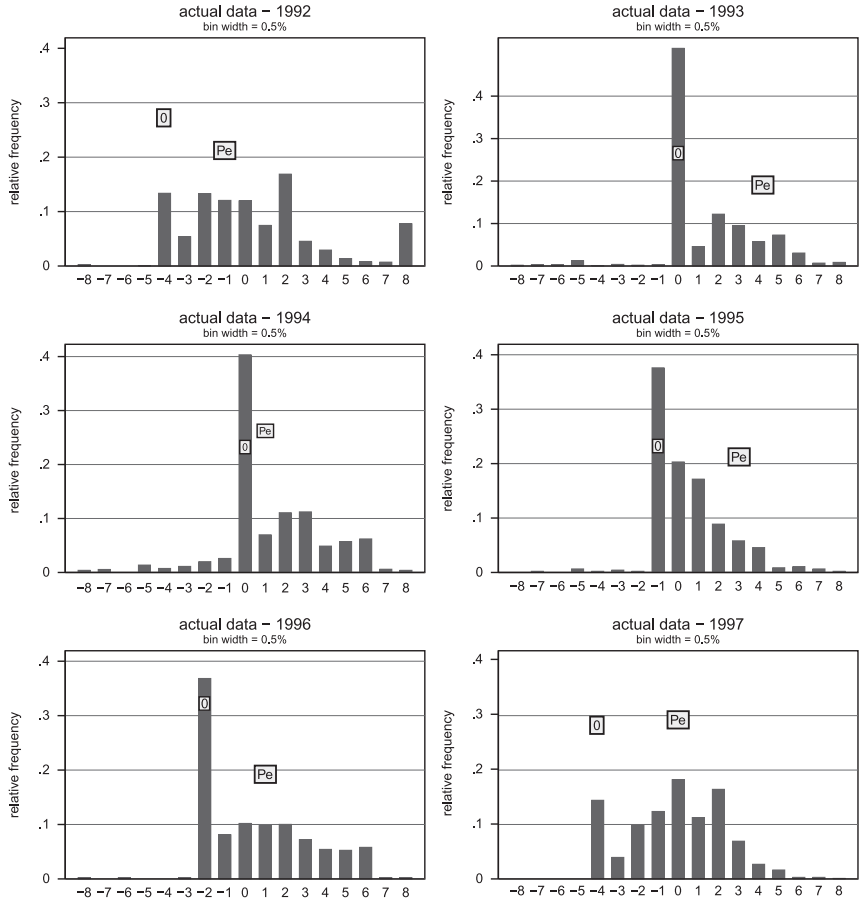

Fig. 3. Standardised (Median-Centred) Relative Frequency Histograms.

probability mass that falls into those intervals (height of bins) to the amount in the corresponding bin of the notional. Bin width selection is driven by the nature of the data and the complexity of distortions that might be involved over intervals.

The bin heights can be formally defined as

$$P_{jt} \equiv \Pr(\eta_{j,t} \leq \dot{w}_{ti} < \eta_{j+1,t}) = \Pr(\dot{w}_{ti} \in \mathcal{B}_{jt}) \tag{1}$$

where $\dot{w}_{ti}$ is the ith observation in year t, and $\mathcal{B}_{jt} \equiv [\eta_{j,t}, \eta_{j+1,t})$ is the jth bin of the probability histogram in year t.

In our earlier work, we used relative frequency as the estimator of the height of bins

$$\hat{P}_{jt} \equiv \sum_{i=1}^{n} \frac{I(\eta_{j,t} \leq \dot{w}_{ti} < \eta_{j+1,t})}{n} = \sum_{i=1}^{n} \frac{I(\dot{w}_{ti} \in \mathcal{B}_{jt})}{n} \tag{2}$$

where $I(\cdot)$ an indicator function. This estimator, which could be motivated from the relative frequency definition of probability, is unbiased as well as consistent. However, in the non-parametric density estimation literature, this estimator is believed to suffer from certain problems. In particular, that it gives non-smooth estimates, that, in addition, depend critically on how the bins are defined, both with respect to their width and location. This is the consequence of the estimator under-smoothing the data.[3]

As a robustness check, we also consider an alternative approach to estimating probability histograms that, in theory, overcomes these problems. It is based on kernel CDF estimation. To motivate the new estimator, we re-write Eq. (1) as follows:

$$P_{jt} = F_t(\eta_{j+1,t}) - F_t(\eta_{j,t}) \tag{3}$$

where $F_t(\cdot)$ is the CDF for the data in year t. Then, we get an estimator for the bin heights by plugging-in some estimator of the CDF in Eq. (3). The CDF estimator we consider is based on the kernel estimator of the corresponding PDF.[4] Substituting $f_t(\cdot)$ in the expression that links the CDF with the PDF by its Kernel estimator, we get

$$\begin{aligned} \hat{F}_t(\dot{w}) &= \int_{-\infty}^{\dot{w}} \hat{f}_t(u)\, du \\ &= \int_{-\infty}^{\dot{w}} \left[\frac{1}{hn} \sum_{i=1}^{n} K\left(\frac{u - \dot{w}_{ti}}{h} \right) \right] du = \frac{1}{n} \sum_{i=1}^{n} G\left(\frac{\dot{w} - \dot{w}_{ti}}{h} \right) \end{aligned} \tag{4}$$

where the function $G(\cdot)$ is the integral of the kernel function $K(\cdot)$. The resulting bin height estimator is then given by

$$\hat{P}_{jt} = \frac{1}{n} \sum_{i=1}^{n} \left[G\left(\frac{\eta_{j+1,t} - \dot{w}_{ti}}{h} \right) - G\left(\frac{\eta_{j,t} - \dot{w}_{ti}}{h} \right) \right] \tag{5}$$

This is consistent, but only asymptotically unbiased. Furthermore, it coincides with the relative frequency estimator when the bandwidth is set equal to zero.

To apply this estimator, we need to choose the type of kernel function and the bandwidth h. For the work described here we have used the Epanechnikov kernel and the least squares cross validation method to choose the optimal bandwidth.[5] This approach provides alternative estimates of the stage 1 histogram heights and we check the robustness of our stage 2 results using this alternative method of construction.

2.3. Tailoring the Model to Inflation Sub-Periods

The data we use are drawn from three fairly distinct periods: High inflation (1977–1982), moderate inflation (1983–1991) and low inflation (1992–1997). These should be characterised by different types of rigidity and present modelling challenges that are explored in detail in Section 5.

3. DATA FEATURES

The data used in this study is derived from 10,945 collective bargaining agreements reached in all of the industries and regions of Canada between 1996 and 1999.[6] These are legally binding agreements, records for which are kept by Human Resources Development Canada (as it was known at the time the data was released to us), or HRDC. These agreements cover bargaining units involving 200 to nearly 80,000 employees, or approximately 11% of the working population of Canada in the mid-year of 1989. They are derived from both the the private and the public sector, and their duration ranges from a few months to several years. Because reporting requirements apply, this information is thought to be very accurate. The dataset used in the empirical work below contains one observation from each of the 10,945 contracts which provides the rate of growth of the basic nominal wage rate. This growth rate refers to the total wage adjustment in the contract, including increases occasioned by the cost of living allowance (COLA) clause. It should be noted, however, that, because the incidence and intensity of COLA clauses is limited throughout the observation period, the results we obtain are similar to those that could be obtained based on non-contingent wage adjustment alone. The observation for each contract is the growth rate of the total nominal-wage adjustment over the whole of the life of the contract, calculated at *annual* rates and is allocated to the year that the contract became effective.

The data from HRDC is supplemented with information from Statistics Canada on the consumer price index (CPI) inflation and an estimate of the mean anticipated inflation ($\widehat{\dot{P}^e}$) for each year.[7] These, along with the median value of the WGD for each year, appear in Table 1. From the CPI figures, the observation period can be divided into three consecutive periods of inflation: 1977–1983 is a high-inflation period with average inflation of 9.58%; 1984–1992 is a medium-inflation period with average inflation of 4.67% and 1993–1997 is a low-inflation period with average inflation of 1.46%. There is obviously a positive relationship between the yearly $\dot{\text{CPI}}$, $\widehat{\dot{P}^e}$ and the median of the realised WGD. There is also a positive relationship between the level of realised inflation, the spread of anticipated inflation and the spread of the WGDs. The latter is visually evident in Figs. 1–3, where the annual histograms of the data are shown.[8] It should be noted that, within each inflation sub-period, the spread of the WGD is relatively constant.

Table 1. Descriptive Statistics.

Year	Obs.	Med($\dot{w}_t$)	$\dot{\text{CPI}}$	$\widehat{\dot{P}^e}$	$\widehat{\text{Var}}(\widehat{\dot{P}^e})$
1977	226	8.20	7.55	7.22	0.4217
1978	673	7.43	8.01	8.42	0.4037
1979	569	10.11	8.95	8.45	0.3680
1980	520	11.95	9.13	9.28	0.3307
1981	450	13.10	10.16	11.66	0.3120
1982	562	10.69	12.43	10.43	0.2737
1983	643	5.00	10.80	6.05	0.3342
1984	676	4.00	5.86	4.50	0.3357
1985	519	4.04	4.30	3.81	0.3185
1986	551	4.10	3.96	4.08	0.2682
1987	557	3.83	4.18	4.37	0.2311
1988	556	4.89	4.34	3.97	0.1919
1989	493	5.22	4.05	4.83	0.1236
1990	547	5.77	4.99	4.55	0.1282
1991	530	4.19	4.76	5.91	0.4946
1992	632	2.00	5.62	1.49	0.7411
1993	516	0.00	1.49	2.00	0.4902
1994	471	0.00	1.86	0.50	0.4740
1995	460	0.68	0.16	2.24	0.4620
1996	448	0.87	2.16	1.43	0.4299
1997	346	1.87	1.62	1.95	0.3528
Total	10,945				

Only 102 (or 0.9%) of the 10,945 contracts in the sample involve nominal wage cuts, while a substantial number (1,142 or 10.4%) show a wage freeze; jointly, these figures could indicate evidence in favour of DNWR. The wage freezes are particularly pronounced during the low-inflation years; for each of the years 1993–1996 the proportion of contracts with a wage freeze was above 35%, peaking at 51.0% in 1993. On the other hand, 6,045 (or 55.2%) of the contracts exhibit negative real wage growth, while 4,801 of them had at the same time positive nominal wage growth. These indications of real wage flexibility must be interpreted with care since they do not rule out DRWR, as has been pointed out. The number of contracts that had exactly zero real wage growth is just 1, and the remaining 4,899 (or 44.8%) contracts showed both nominal and real wage increase. The econometric approach used to examine DWR is now described.

4. EMPIRICAL SPECIFICATION

A detailed description of the econometric approach followed in our work appears in Christofides and Nearchou (2007). The basic idea is to test hypotheses about the shape of the actual WGD in terms of the heights of the bins of the corresponding probability histogram. We first proceed to express the actual WGDs for each year into histograms, which are then estimated non-parametrically. The resulting estimates are then used, in a second stage, in econometric estimation, where we estimate jointly the notional distribution and the distortions due to DWR.

4.1. Outline of the Testing Methodology

Testing for the presence of either type of DWR takes the form of testing hypotheses about the shape of the WGDs. Our approach is to describe the WGD with a probability histogram. Hence, the testing of hypotheses about the shape of WGDs takes the form of testing hypotheses about the height of the bins of the corresponding probability histogram.

The probability histogram for the WGD of year t could be defined as the collection of probabilities $\{P_{jt}\}_{j=-J}^{J}$, where j is the bin index. Given that our analysis focuses on the shape of the WGDs but not their location, j is defined to indicate the position of the bins relative to each other, rather than the real line. In particular, the bin indexed by $j = 0$ contains the median of the actual WGD, bins indexed by a negative j lie $|j|$ positions to the left of

the median bin and bins indexed by a positive j lie j positions to its right. Furthermore, the bins of the histogram are defined such that the median is located at the centre of the 'median' bin. We describe the probability histograms defined in this way as 'standardised', using median centering.

In order to formulate the relevant tests, we parameterise P_{jt} under the hypotheses of no rigidity and DWR, respectively, by

$$P_{jt} = \begin{cases} p^N(z_{jt}^N; b_j^N), & \text{if } H_0 \text{ is true} \\ p^R(z_{jt}^R; b_j^R), & \text{if } H_1 \text{ is true} \end{cases} \tag{6}$$

where $p^N(\cdot)$ is the function of a vector of observables z_{jt}^N that gives the height of the jth bin of the probability histogram of the notional distribution in year t, $p^R(\cdot)$ the function of observables z_{jt}^R that gives the height of the corresponding bin of the probability histogram of the rigidity-contaminated distribution in the same year, and b_j^N and b_j^R the corresponding vectors of parameters. Typically both z_{jt}^N and z_{jt}^R will contain dummy variables that will be functions of j and will indicate the relative position of bin j in the probability histogram; they may also contain additional variables that capture characteristics of the year t, while z_{jt}^R will additionally contain variables that indicate the position of bin j relative to the position of the bins containing the values taken by the rigidity bounds in the population. These variables will be functions of both j and the corresponding indices of the bins that contain the point zero (i.e. the rigidity bound for DNWR), and the anticipated inflation values (i.e. the rigidity bounds for DRWR).

With this formulation, we could test hypotheses about DWR by estimating the unrestricted model (with rigidity), and, subsequently, testing hypotheses, about the parameter vector b_j^R, that imply that the unrestricted model coincides with the restricted (rigidity-free). We implement this approach in two stages.

In stage 1, the probability histogram describing the distribution underlying the observed wage growth data for each year in the sample is estimated non-parametrically. In stage 2, for each j, using the set of T estimates of the height of bin j from all years, i.e. $\{\hat{p}_{jt}\}_{t=1,\ldots,T}$, as the set of 'observations' on $\hat{P}_{jt}$,[9] we estimate the regression of $\hat{P}_{jt}$ on the vector of observables z_{jt}^R. When the estimator $\hat{P}_{jt}$ is unbiased, the regression function will coincide with $p^R(z_{jt}^R; b_j^R)$.[10] Therefore, the estimation of this equation would give estimates of the parameter vector b_j^R and its variance–covariance matrix, enabling us to test a number of restrictions related to DWR.

In practice, the regression equations corresponding to all bin heights are estimated jointly since this is typically more efficient.[11]

4.2. Parameterisation of Probability Histograms

In this section, we describe the most general specification of the model for the bin heights, which is estimated with the full sample. For the sub-periods, we trim this specification in order to accommodate for the special features of these periods.

Our chosen parameterisation for the heights of the bins of the probability histograms under the null hypothesis (i.e. for the notional[12] distribution), is the following

$$\begin{aligned} p^N(z_{jt}^N; b_j^N) &= \beta_{1|j|} + \beta_{2|j|} \times up_{jt} + (\beta_{3|j|} + \beta_{4|j|} \times up_{jt}) \times m_t, \quad j \neq 0 \\ &= \beta_{10} + \beta_{30} \times m_t, \quad j = 0 \end{aligned} \tag{7}$$

where m_t denotes the median of the actual-wage-growth data in year t, up_{jt} is a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ lies to the right of the bin containing the median ($j > 0$), and the βs are coefficients to be estimated. With this parameterisation the $2J + 1$ probability bins in each histogram can have different height from each other, therefore, the notional distribution is not restricted to have any particular shape or to be symmetric. Furthermore, by making the bin height to be a linear function of the location of the actual WGD, and therefore of the location of the notional distribution itself, we allow for the shape of the notional distribution to vary with its location. For example, suppose that the notional distribution is symmetric around the bin containing m_t and, further, that its spread increases as its centre moves to higher values.[13] Then $\beta_{2|j|}$ and $\beta_{4|j|}$ will be equal to zero due to the symmetry assumption, $\beta_{1|j|}$ will be non-negative, and $\beta_{3|j|}$ will be negative for the bins in the middle of the distribution, that is for small $|j|$, and positive for the bins that lie to the tails of the distribution, that is for large $|j|$. Alternatively, if we allow $\beta_{4|j|}$ to be non-zero for some values of j, then the skewness of the notional distribution will also vary with the location.[14]

In order to test for the presence of both types of rigidity, the parameterisation of the probability histogram under the alternative hypothesis should reflect the distortions due to the presence of both. We assume that

$$p^R(z_{jt}^R; b_j^R) = p^N(z_{jt}^N; b_j^N) + D^u(z_{jt}^u; \mu) + D^n(z_{jt}^n; \gamma) + D^r(z_{jt}^r; \delta), \text{for } R = nr \tag{8}$$

where $D^n(z^n_{jt};\gamma)$ is defined to be the difference between the height of the jth bin of the rigidity-contaminated probability histogram and the height of the corresponding bin of the notional probability histogram in year t that is due to the presence of DNWR, and $D^r(z^r_{jt};\delta)$ the corresponding difference that is due to the presence of DRWR. We also allow for distortions due to the presence of menu costs, captured by the term $D^u(z^u_{jt};\mu)$.

For distortions due to DNWR, we write

$$D^n(z^n_{jt};\gamma) = (\gamma_1 + \gamma_2 \times m_t) \times d0_{jt} + (\gamma_3 + \gamma_4 \times m_t) \times dn_{jt} + \gamma_5 \times dz1_{jt} \quad (9)$$

where $d0_{jt}$ is a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ contains the point zero, dn_{jt} a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ is to the left of the bin containing the point zero, and $dz1_{jt}$ a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ is the first bin to the right of the bin that contains the point zero. With the inclusion of the first term, we can capture the distortion that applies to the bin that contains zero nominal wage growth, and, with the second term, the distortion that applies to each one of the bins that contain negative values of wage growth. In particular, γ_1 accounts for the distortion associated with the bin that contains zero nominal wage growth and γ_3 the distortion associated with the bins that lie to the left of this bin in the special case where the centre of the notional distribution, which we proxy by m_t, is located at the point zero (i.e. $m_t = 0$). In that case, and, in the presence of DNWR, we would expect γ_1 to be positive, signifying the concentration of probability mass surplus in the zero nominal wage growth bin, and γ_3 negative, signifying the loss of probability mass from the bins that contain negative values of notional wage growth. When the centre of the notional distribution is located further to the right ($m_t > 0$), a smaller part of the left tail of the notional distribution lies below zero, that is the proportion of notional wage cuts falls, and, therefore the proportion of notional wage changes that become wage freezes due to DNWR is expected to fall. In that case, γ_2 must be negative, signifying the reduction in the probability mass surplus in the zero nominal wage growth bin, while γ_4 could be either positive or negative or zero, as the amount of mass deficit from each bin containing negative values could change in any direction relative to its level at $m_t = 0$. The inclusion of the last term enables us to test the hypothesis that, apart from shifting mass to the point of zero nominal wage growth, the presence of DNWR could also induce a shift of mass beyond the point zero, towards small positive values (in that case, $\gamma_5 > 0$) – see Holden (1989, 1998, 2004) and Cramton and Tracy (1992).

The distortion in the height of the probability bar of bin $\mathcal{B}_{jt}$ due to DRWR is assumed to be given by

$$D^r(z^r_{jt};\delta) = \delta_{1k} + \delta_{2k} \times J^P_t, \quad k = j - J^P_t, \ k_{\min} \le k \le k_{\max} \tag{10}$$

$$= \sum_{v=k_{\min}}^{k_{\max}} (\delta_{1v} + \delta_{2v} \times J^P_t) \times dp_{v,jt} \tag{11}$$

where J^P_t is the value of the index of the bin in year t that contains $\widehat{\dot{P}^e}$ (estimated mean of AID), k is the distance between bin $\mathcal{B}_{jt}$ and that bin,[15] and $dp_{v,jt}$ are dummy variables indicating whether bin $\mathcal{B}_{jt}$ is located k positions from the bin that contains the centre of the AID in year t,

$$dp_{v,jt} = \begin{cases} 1, & \text{if } v = k(= j - J^P_t) \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

With this specification, we allow for the size of the distortions to differ according to the location of the bin in the support of the AID (through the indexing by k), and its location in the support of the notional WGD (through the dependence on J^P_t). In the presence of DRWR, the δ_{1k} coefficients, which account for the distortion when the centre of the AID is located in the same bin as the median of the actual-wage-growth distribution ($J^P_t = 0$), are expected to be positive for the largest (and positive) values of k and negative for the smallest (and negative) values of k, signifying the shift of probability mass towards the right end of the support of the AID. When J^P_t takes different values, the values of the δ_{2k} coefficients must be such that the distortions ($\delta_{1k} + \delta_{2k} \times J^P_t$) are qualitatively similar to the case where $J^P_t = 0$, however no specific statements can be made about their sign or size unless specific assumptions are made about the nature of the joint distribution of the notional-wage growth and anticipated inflation, and the rigidity mechanism.

Finally, the effect of menu costs is parameterised as follows

$$D^u(z^u_{jt};\mu) = \mu \times dnp1_{jt} \tag{13}$$

where $dnp1_{jt}$ is a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ is either one position to the left or to the right of the bin that contains the point zero. Therefore, we allow for a symmetric loss of mass ($\mu < 0$) around and close to zero.

For the identification of the parameters of the model, it is required that each type of rigidity distort different parts of the WGD at least for some of the years in the sample. In this way, there will be sufficient variation in the

dummy variables that indicate the bins that are affected by the distortions, so that these will not be collinear with the dummy variables that indicate the position of the bins in the notional probability histogram. This identification strategy is most relevant to the whole-sample period where a rich inflation experience can be found. Where sub-periods are concerned, it is important to keep in mind the unique features of the period and to modify the identification strategy.

4.3. Estimation

For the estimation of the probability histograms in stage 1, we consider two alternative estimators; the relative frequency estimator, described by Eq. (2), and the kernel-based estimator, described by Eq. (5).

Regarding the estimation in stage 2, the exact algebraic expression for the covariance between any pair of estimators that correspond to bins from the same or different probability histograms was derived, for the relative frequency case of stage 1, in Christofides and Nearchou (2007). This allows for the preferred estimator (FGLS) to be implemented but we also reported results based on ordinary least squares (OLS) along with the corrected standard errors (corrected OLS). The relative frequency and Kernel approaches produce stage 1 data that are very similar indeed. As a result, the stage 2 parameter estimates are also very similar and, because of space limitations, they are not reported. Details are available on request.

5. RESULTS

5.1. Whole-Sample Results

As the first step, we implemented the model in Christofides and Nearchou (2007) but using the median-centered data discussed above. The results obtained are so similar that, in the interests of economy, are not presented here. In what follows, we always, therefore, use the median-centered data.

A natural next question is whether improvements to the specification of our earlier work can be achieved, given the new median-centered data. Small improvements are possible. In Table 2, we present results for the whole sample, median-centered data, based on FGLS and Corrected OLS. We have attempted to achieve parsimony in the specification for the effects of DRWR in the area to the right of the bin containing $\widehat{\dot{P}^e}$ because our

Table 2. Estimation Results: Full Sample.

Parameter	FGLS		Corrected OLS	
	Estimate	SE	Estimate	SE
β_{10}	0.3571*	0.0091	0.3055*	0.0110
β_{11}	0.1060*	0.0062	0.1963*	0.0103
β_{12}	0.0645*	0.0055	0.0883*	0.0094
β_{13}	0.0508*	0.0053	0.0788*	0.0082
β_{14}	0.0374*	0.0051	0.0513*	0.0073
β_{15}	0.0132*	0.0049	0.0535*	0.0062
β_{16}	0.0408*	0.0082	0.0518*	0.0048
β_{17}	0.0247*	0.0065	0.0538*	0.0039
β_{18}	0.0178*	0.0049	0.0560*	0.0039
β_{21}	0.1558*	0.0098	0.0345*	0.0126
β_{22}	0.0133***	0.0080	−0.0101	0.0112
β_{23}	−0.0297*	0.0065	−0.0443*	0.0094
β_{24}	−0.0359*	0.0053	−0.0414*	0.0077
β_{25}	−0.0130**	0.0052	−0.0542*	0.0063
β_{26}	−0.0395*	0.0085	−0.0543*	0.0050
β_{27}	−0.0244*	0.0065	−0.0557*	0.0040
β_{28}	−0.0158*	0.0054	−0.0567*	0.0040
β_{30}	−0.0223*	0.0010	−0.0152*	0.0012
β_{31}	0.0061*	0.0008	−0.0039*	0.0011
β_{32}	0.0085*	0.0007	0.0037*	0.0009
β_{33}	0.0029*	0.0005	−0.0013***	0.0007
β_{34}	0.0009**	0.0004	−0.0014**	0.0006
β_{35}	0.0015*	0.0004	−0.0019*	0.0006
β_{36}	−0.0018**	0.0007	−0.0032*	0.0004
β_{37}	−0.0009***	0.0005	−0.0035*	0.0004
β_{38}	−0.0006	0.0005	−0.0038*	0.0003
β_{41}	−0.0186*	0.0014	−0.0063*	0.0016
β_{42}	−0.0060*	0.0012	−0.0013	0.0014
β_{43}	0.0004	0.0008	0.0028*	0.0009
β_{44}	0.0025*	0.0006	0.0041*	0.0008
β_{45}	0.0003	0.0005	0.0045*	0.0007
β_{46}	0.0030*	0.0008	0.0055*	0.0006
β_{47}	0.0015*	0.0005	0.0048*	0.0005
β_{48}	0.0012**	0.0006	0.0046*	0.0004
μ	−0.0134*	0.0026	−0.0221*	0.0016
γ_1	0.0988*	0.0051	0.1615*	0.0102
γ_2	−0.0154*	0.0010	−0.0293*	0.0019
γ_3	−0.0158*	0.0017	−0.0603*	0.0038
γ_4	0.0002	0.0004	0.0066*	0.0004
γ_5	0.0128*	0.0035	0.0092***	0.0050
δ_{-18}	−0.0088*	0.0030	−0.0030*	0.0010
δ_{-17}	0.0049***	0.0026	0.0021	0.0018

Table 2. (*Continued*)

Parameter	FGLS		Corrected OLS	
	Estimate	SE	Estimate	SE
δ_{-16}	−0.0117*	0.0036	0.0010	0.0037
δ_{-15}	−0.0168*	0.0041	−0.0034	0.0053
δ_{-14}	−0.0268*	0.0044	−0.0186*	0.0067
δ_{-13}	−0.0333*	0.0047	−0.0167**	0.0078
δ_{-12}	−0.0115**	0.0052	−0.0112	0.0078
δ_{-11}	0.0175*	0.0057	−0.0087	0.0073
δ_{10}	0.0398*	0.0054	0.0237*	0.0065
δ_{11}	0.0179*	0.0045	0.0177*	0.0053
δ_{-28}	0.0076	0.0058	0.0051*	0.0019
δ_{-27}	0.0009	0.0015	0.0006	0.0020
δ_{-26}	−0.0023***	0.0012	0.0046**	0.0022
δ_{-25}	−0.0023**	0.0012	−0.0004	0.0019
δ_{-24}	−0.0045*	0.0012	0.0005	0.0018
δ_{-23}	−0.0075*	0.0014	−0.0015	0.0021
δ_{-22}	0.0068*	0.0017	0.0064**	0.0028
δ_{-21}	0.0157*	0.0024	0.0030	0.0036
δ_{20}	0.0037	0.0030	−0.0075**	0.0037
Obs.	357		357	

*Significance level at 1%.
**Significance level at 5%.
***Significance level at 10%.
Note: The parameters $\beta_{1|j|}$ and $\beta_{3|j|}$ refer to the symmetric part of the notional distribution, while the parameter $\beta_{2|j|}$ and $\beta_{4|j|}$ allow this distribution to be non-symmetric. The parameter μ refers to the menu-cost behaviour. The γs capture DNWR behaviour, with $\gamma_1 + \gamma_2 m_t$ measuring the spike at zero and $\gamma_3 + \gamma_4 m_t$ the deficit in the bins that contain negative values, where m_t is the median of the actual WGD from period t. The δs capture DRWR: when the median WGD bin also contains the expected inflation rate ($J_t^P = 0$), then δ_{10} measures the extra mass due to DRWR in that bin, with parameters $\{\delta_{-11}, \ldots, \delta_{-18}\}$ measuring distortions to bins that lie to its left, and $\{\delta_{11}\}$ the distortion to the bin that lies to its right. Please see the text for a more complete explanation.

variance estimates in column 5, Table 1 suggests that the AID is quite tight and because we want to maintain some degree of comparability with specifications for the sub-periods. Note that, since DRWR could shift mass from below the minimum point in the support of the AID, similar parsimony to the left of the bin containing $\widehat{P^e}$ is not desirable.

It is clear that all the qualitative features of the earlier paper are present. DNWR is clearly present. When the median of the WGD is zero, this type of rigidity accounts for an accumulation of nearly 9.88 percentage points of

mass at the point zero and for a reduction of 1.58 points of mass in each of the bins involving negative wage growth (FGLS). The spike at zero becomes smaller as the median increases; if it were to increase to 4% (the approximate value of the median for the sample as a whole), the additional spike at zero would be 3.72 percentage points ($9.88 - 1.54 \times 4$). It will be seen below that these whole-sample estimates of DNWR average substantially higher effects during the low-inflation period with lower effects in the other sub-periods. The distortions due to DRWR are well-defined and in line with our expectations: When the 'median' bin also contains $\widehat{\dot{P}^e}$, that bin attracts 3.98 percentage points of additional mass and its adjacent bins about 1.7 percentage points of additional mass (FGLS results), as an approximately equal mass gets shifted from bins further to the left to the bins mentioned above. Again, these are average effects for the whole sample. These results are modified by the further interactions and menu-cost effects that are allowed for and the rigidity-contaminated and notional distributions (FGLS) appear more clearly in Fig. 4, for selected years.

The apparent ability of the model to pick up the distortions occasioned by DWR is, to a large extent, due to the rich inflation experience present in the whole sample. It is, now, of interest to see how this model may be applied to the sub-periods. Since these involve fewer observations, it will be necessary to both simplify the model but also to adapt it to suit the needs and challenges of the sub-periods. We simplify by generally omitting consideration of menu-cost behaviour ($\mu = 0$), of the effect discussed by Holden (1989, 1998, 2004) and Cramton and Tracy (1992) ($\gamma_5 = 0$), of changes in the notional distribution as the median changes ($\beta_{3|j|} = \beta_{4|j|} = 0,\ \forall j$), and of changes in the DRWR-induced distortions that may occur as the actual WGD and AID shift around ($\delta_{2k} = 0,\ \forall k$). The latter assumptions are justified by the fact that these shifts are necessarily more limited within the sub-periods. The changes reduce substantially the number of parameters that must be estimated during the sub-periods.

5.2. Sub-Samples

5.2.1. High Inflation (1977–1982)

During the period 1977–1982,[16] the WGD for the individual years does not often involve negative wage change and DNWR may not be relevant. Although DRWR distortions are the only ones that should be expected during high-inflation periods, their identification in practice may be difficult if these periods are short in duration (yielding a small number of yearly

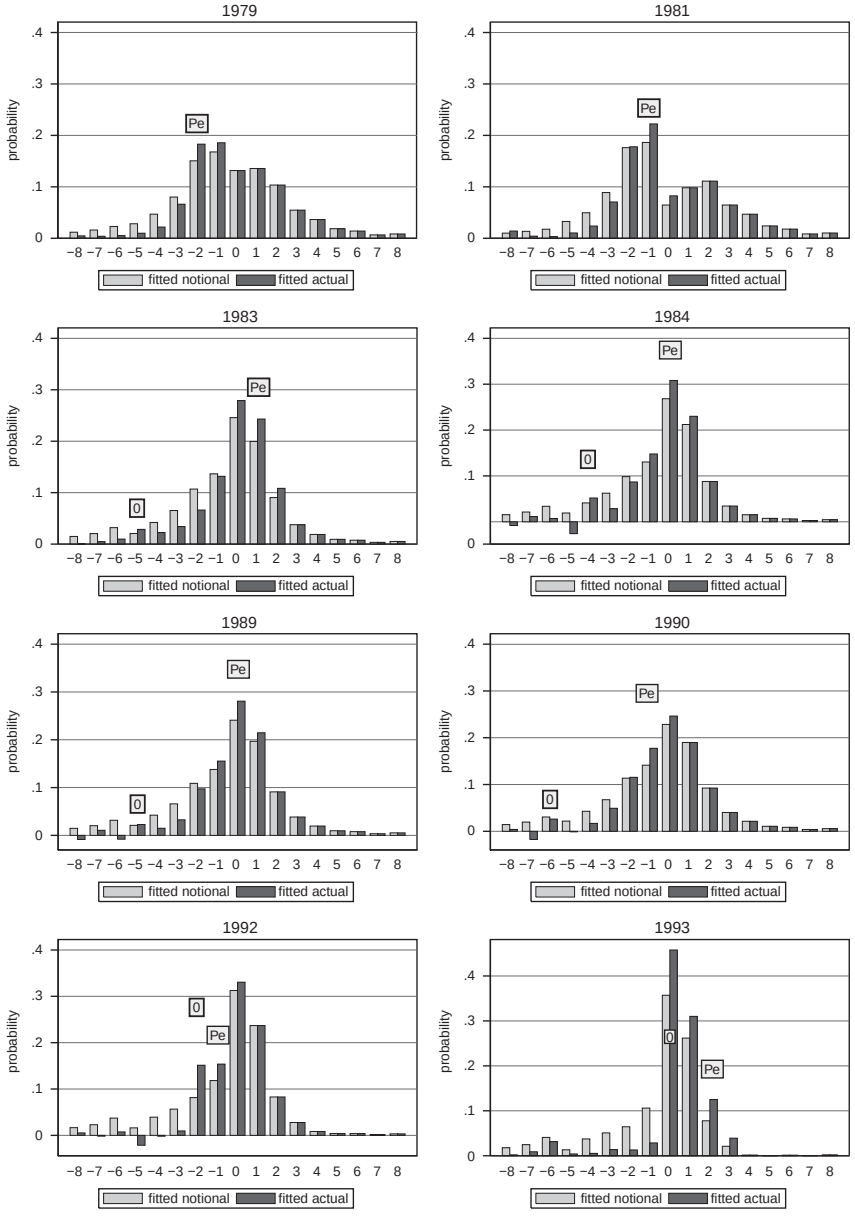

Fig. 4. Notional Versus Actual Nominal WGDs (Fitted Values): Full Sample, FGLS Results in Table 2 (Diagrams for Selected Years).

samples) and the difference between $\widehat{\dot{P}^e}$ and the median of the WGD is not sufficiently rich. To the extent that any distortions can be identified, these are, most likely, due to DRWR. Table 3 presents FGLS and corrected OLS results for a version of the model that was simplified as described in the previous section. The parameter γ_1 for DNWR is not significantly different from zero, as one would expect. While the shift of mass towards the bin

Table 3. Estimation Results: High-Inflation Period.

Parameter	FGLS		Corrected OLS	
	Estimate	SE	Estimate	SE
β_{10}	0.1396*	0.0068	0.1419*	0.0074
β_{11}	0.1613*	0.0072	0.1615*	0.0082
β_{12}	0.1342*	0.0066	0.1348*	0.0081
β_{13}	0.0660*	0.0049	0.0647*	0.0066
β_{14}	0.0293*	0.0036	0.0310*	0.0055
β_{15}	0.0114*	0.0027	0.0262*	0.0053
β_{16}	0.0089*	0.0025	0.0089**	0.0044
β_{17}	0.0052**	0.0021	0.0041	0.0042
β_{18}	0.0055***	0.0028	0.0009	0.0031
β_{21}	−0.0311*	0.0096	−0.0426*	0.0103
β_{22}	−0.0180**	0.0090	−0.0240**	0.0103
β_{23}	−0.0035	0.0066	−0.0071	0.0083
β_{24}	0.0145*	0.0052	0.0109	0.0069
β_{25}	0.0216*	0.0041	0.0067	0.0064
β_{26}	0.0168*	0.0037	0.0161*	0.0055
β_{27}	0.0062**	0.0028	0.0106**	0.0049
β_{28}	0.0040	0.0033	0.0082**	0.0037
γ_1	−0.0006	0.0033	−0.0017	0.0045
δ_{-18}	−0.0019	0.0033	0.0062	0.0052
δ_{-17}	0.0065***	0.0036	0.0025	0.0038
δ_{-16}	−0.0010	0.0023	0.0021	0.0043
δ_{-15}	−0.0004	0.0024	0.0096**	0.0046
δ_{-14}	−0.0031	0.0025	−0.0093***	0.0050
δ_{-13}	−0.0046	0.0031	−0.0158*	0.0057
δ_{-12}	−0.0047	0.0036	0.0011	0.0070
δ_{-11}	−0.0047	0.0057	−0.0049	0.0079
δ_{10}	0.0086	0.0073	0.0068	0.0084
δ_{11}	0.0011	0.0074	0.0039	0.0083
N	102		102	

*Significance level at 1%.
**Significance level at 5%.
***Significance level at 10%.

Table 4. Estimation Results: High-Inflation Period.

Parameter	FGLS		Corrected OLS	
	Estimate	SE	Estimate	SE
β_{10}	0.1435*	0.0067	0.1368*	0.0072
β_{11}	0.1486*	0.0046	0.1353*	0.0052
β_{12}	0.1302*	0.0043	0.1189*	0.0048
β_{13}	0.0689*	0.0030	0.0586*	0.0034
β_{14}	0.0396*	0.0023	0.0362*	0.0027
β_{15}	0.0222*	0.0018	0.0313*	0.0025
β_{16}	0.0184*	0.0017	0.0203*	0.0021
β_{17}	0.0108*	0.0013	0.0134*	0.0018
β_{18}	0.0100*	0.0014	0.0086*	0.0014
γ_1	−0.0010	0.0032	−0.0013	0.0045
δ_{-18}	−0.0058***	0.0031	−0.0025	0.0042
δ_{-17}	0.0036	0.0033	−0.0067*	0.0021
δ_{-16}	−0.0065*	0.0017	−0.0065**	0.0025
δ_{-15}	−0.0058*	0.0021	0.0013	0.0029
δ_{-14}	−0.0113*	0.0019	−0.0143*	0.0027
δ_{-13}	−0.0138*	0.0025	−0.0166*	0.0037
δ_{-12}	−0.0133*	0.0030	0.0084	0.0053
δ_{-11}	−0.0065	0.0051	0.0058	0.0069
δ_{10}	0.0080	0.0071	0.0173**	0.0080
δ_{11}	−0.0022	0.0073	0.0086	0.0081
N	102		102	

*Significance level at 1%.
**Significance level at 5%.
***Significance level at 10%.
Note: Symmetry ($\beta_{21} = \ldots = \beta_{28} = 0$) is imposed.

containing $\widehat{\dot{P}^e}$ is, to an extent, apparent, these distortions (δ_{11} to δ_{-18}) are not generally significant. This may be due to the limited number of observations involving diverse points on the WGD. We simplify the model further by imposing symmetry on the notional distribution, an assumption that has been used in several earlier papers. Table 4 shows that significant shifts in mass occur from several points to the left of the bin containing $\widehat{\dot{P}^e}$. The gains in the bin containing $\widehat{\dot{P}^e}$ are statistically significant in the case of the corrected OLS results, with a gain in the bin containing $\widehat{\dot{P}^e}$ equal to 1.73 percentage points. Fig. 5 plots the estimated (FGLS) probability histograms for the notional and actual WGDs based on Table 3, and Fig. 6 the corresponding histograms based on Table 4.

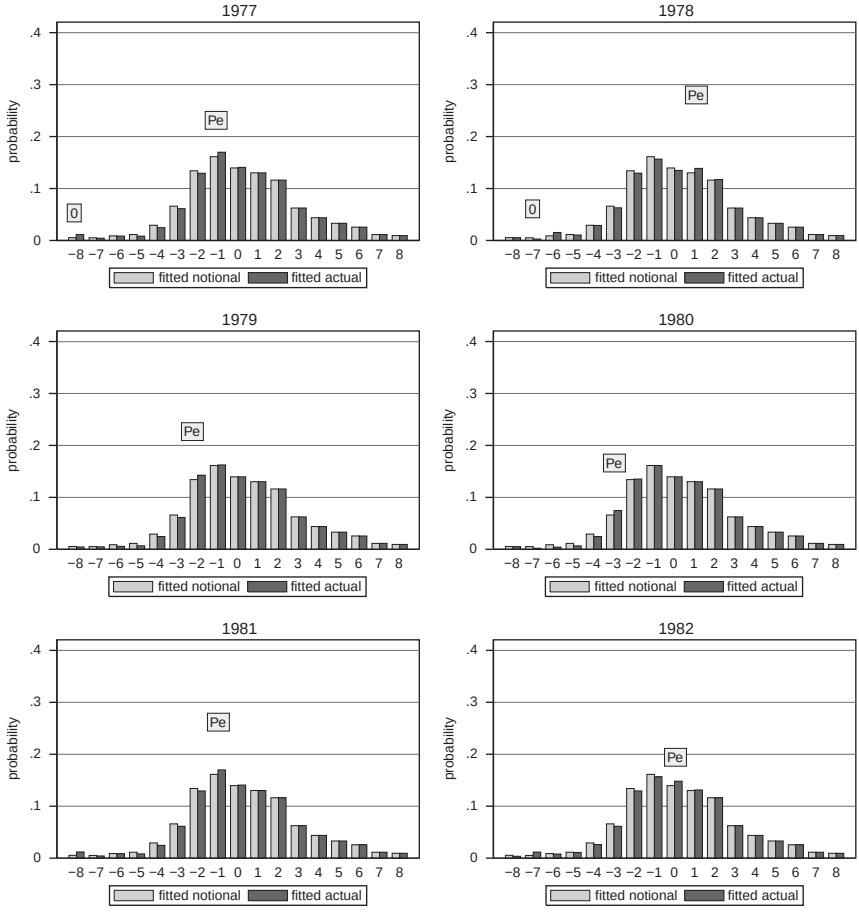

Fig. 5. Notional Versus Actual Nominal WGDs (Fitted Values): High-Inflation Period, FGLS Results in Table 3.

5.2.2. Medium Inflation (1983–1991)

During this period, the WGD in our data extends into the negative orthant in every year of the sample. The mass of the actual WGD which is at, or below, zero is as low as 0.7% in 1988 and as high as 11.2% in 1991. At the same time, $\widehat{\dot{P}^e}$ ranges between 3.81% in 1985 and 6.05% in 1983, substantially above the point zero – see Section 3. Thus, a sizeable distance between the relevant ranges for DNWR and DRWR in the WGD exists and a clear separation and identification of the two processes may be possible.

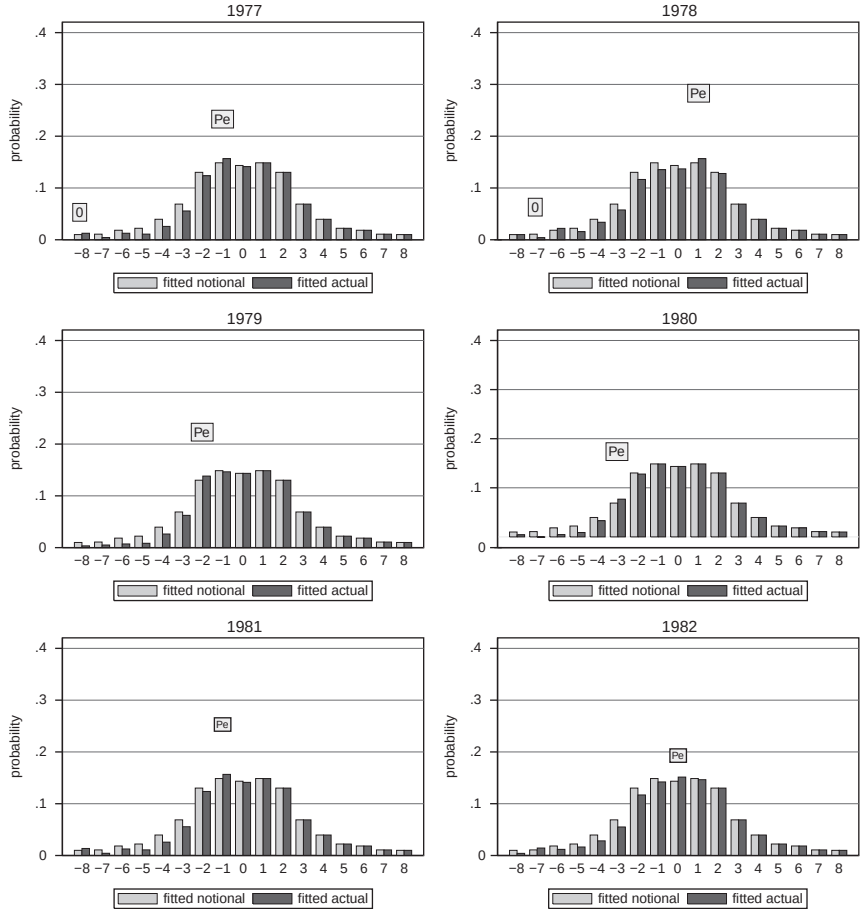

Fig. 6. Notional Versus Actual Nominal WGDs (Fitted Values): High-Inflation Period, FGLS Results in Table 4 (Symmetric National).

Table 5 reports a parsimonious version of the model which retains asymmetry in the notional distribution. This model groups the effects on the second to eighth bin below the bin containing $\widehat{\dot{P}^e}$, thus simplifying the estimation (parameter δ_{-128} refers to this group). The results suggest that both kinds of rigidity are at work. For instance, when the median is 4% (its approximate value in this period), the additional mass in the bin containing the point zero is 2.84 percentage points $(11.36 - 2.13 \times 4)$,

Table 5. Estimation Results: Medium-Inflation Period.

Parameter	FGLS		Corrected OLS	
	Estimate	SE	Estimate	SE
β_{10}	0.2347*	0.0073	0.2529*	0.0086
β_{11}	0.1452*	0.0060	0.1892*	0.0085
β_{12}	0.0992*	0.0045	0.1259*	0.0075
β_{13}	0.0376*	0.0028	0.0861*	0.0074
β_{14}	0.0170*	0.0023	0.0525*	0.0069
β_{15}	0.0090 **	0.0036	0.0352*	0.0067
β_{16}	0.0332*	0.0083	0.0398*	0.0065
β_{21}	0.0606*	0.0089	0.0067	0.0102
β_{22}	−0.0259*	0.0061	−0.0516*	0.0083
β_{23}	−0.0058	0.0038	−0.0533*	0.0078
β_{24}	−0.0024	0.0028	−0.0377*	0.0072
β_{25}	−0.0039	0.0037	−0.0298*	0.0068
β_{26}	−0.0295*	0.0083	−0.0357*	0.0066
γ_1	0.1136*	0.0196	0.1381*	0.0209
γ_2	−0.0213*	0.0045	−0.0211*	0.0042
δ_{-128}	−0.0045*	0.0016	−0.0387*	0.0066
δ_{-11}	0.0309*	0.0068	−0.0042	0.0090
δ_{10}	0.0662*	0.0078	0.0277*	0.0088
δ_{11}	0.0215*	0.0054	0.0197*	0.0069
N	117		117	

*Significance level at 1%.
**Significance level at 5%.
***Significance level at 10%.
Note: DNWR and DRWR effects are allowed for.

while that in the bin containing the median and the point $\widehat{\dot{P}^e}$ is 6.62 percentage points.

It is also interesting to explore the importance of misspecifying the estimating equation, as, for instance, when DRWR is ignored while searching for DNWR, as was done in the early papers in this sub-literature. Tables 6 and 7 show versions of the model which contain only DNWR and only DRWR, respectively. While special cases appear to be successfully implemented, the quantitative effects are somewhat different from those in Table 5, where no exclusion restrictions are imposed. The spike at zero in Table 5 is underestimated by 2.54 percentage points while the shift of mass towards the bin containing the expected inflation rate in Table 6 is overestimated somewhat. These particular results suggest that omitting consideration of DRWR leads to bias and underestimation of the DNWR

Table 6. Estimation Results: Medium-Inflation Period.

Parameter	FGLS		Corrected OLS	
	Estimate	SE	Estimate	SE
β_{10}	0.2743*	0.0058	0.2643*	0.0061
β_{11}	0.1674*	0.0049	0.1806*	0.0054
β_{12}	0.1045*	0.0040	0.0949*	0.0041
β_{13}	0.0337*	0.0024	0.0474*	0.0030
β_{14}	0.0130*	0.0017	0.0138*	0.0021
β_{15}	0.0009	0.0021	−0.0035*	0.0011
β_{16}	0.0232*	0.0075	0.0011*	0.0004
β_{21}	0.0705*	0.0082	0.0298*	0.0088
β_{22}	−0.0206*	0.0056	−0.0131**	0.0059
β_{23}	0.0006	0.0034	−0.0124*	0.0040
β_{24}	0.0019	0.0023	0.0010	0.0027
β_{25}	0.0040	0.0024	0.0089*	0.0015
β_{26}	−0.0196*	0.0075	0.0030*	0.0010
γ_1	0.0882*	0.0171	0.1381*	0.0209
γ_2	−0.0153*	0.0039	−0.0211*	0.0042
N	117		117	

*Significance level at 1%.
**Significance level at 5%.
***Significance level at 10%.
Note: DRWR effects are suppressed.

effects. Fig. 7 plots the estimated (FGLS) probability histograms for the notional and actual WGDs based on Table 5.

5.2.3. Low Inflation (1992–1997)

When the median of the WGD is close to zero (as in 1993 and 1994), the extent to which separate DNWR and DRWR can be identified is unclear. The model must be calibrated to avoid undue overlap between DNWR- and DRWR-dedicated dummy variables. Allowing for too many bins may be inappropriate and it is necessary to also explore the possibility of finer binning. We have, therefore, redesigned the stage 1 data to allow for 0.5 percentage point wide bins so as to have a better chance of capturing the detail between the point zero and the rather low values of anticipated inflation (at most 2.24 in 1995, Table 1). We have also allowed for a more flexible specification for the distortions to the bins with negative values, by replacing the dummy variable *dn* (Eq. (9)) with bin specific dummies $dn\zeta$ ($\zeta = 1, \ldots, 6$), where ζ indicates the position of the bin to the left of the bin

Table 7. Estimation Results: Medium-Inflation Period.

Parameter	FGLS		Corrected OLS	
	Estimate	SE	Estimate	SE
β_{10}	0.2367*	0.0073	0.2529*	0.0086
β_{11}	0.1445*	0.0060	0.1892*	0.0085
β_{12}	0.0972*	0.0045	0.1259*	0.0075
β_{13}	0.0343*	0.0027	0.0861*	0.0074
β_{14}	0.0172*	0.0022	0.0820*	0.0076
β_{15}	0.0046***	0.0026	0.0458*	0.0067
β_{16}	0.0086**	0.0037	0.0416*	0.0065
β_{21}	0.0642*	0.0089	0.0067	0.0102
β_{22}	−0.0227*	0.0061	−0.0516*	0.0083
β_{23}	−0.0020	0.0037	−0.0533*	0.0078
β_{24}	−0.0023	0.0027	−0.0672*	0.0079
β_{25}	−0.0014	0.0028	−0.0404*	0.0068
β_{26}	−0.0050	0.0037	−0.0375*	0.0065
δ_{-128}	−0.0005	0.0014	−0.0387*	0.0066
δ_{-11}	0.0348*	0.0067	−0.0042	0.0090
δ_{10}	0.0693*	0.0078	0.0277*	0.0088
δ_{11}	0.0225*	0.0054	0.0197*	0.0069
N	117		117	

*Significance level at 1%.
**Significance level at 5%.
***Significance level at 10%.
Note: DNWR effects are suppressed.

that contains point zero.[17] Estimates appear in Table 8. The astonishing concentration of mass at the bin containing the point zero is evident in the estimate for the additional height in that bin (this is also the median bin in 1993 and 1994). This bin attracts 36.01 points of additional mass. The DRWR mechanism can also be identified. Mass is shifted from points in the left of, to points near and at the bin containing $\widehat{\dot{P}^e}$. For instance, the extra mass at $\widehat{\dot{P}^e}$ is 1.88 percentage points (FGLS). Table 9 suggests that if, as would have been the case in the early years of this literature, a model were fitted for DNWR only, the results (Table 9) would credit to DNWR concentration of mass that in the more general specification of Table 8 belongs to DRWR, thereby overestimating DNWR by about two percentage points. Thus, suppressing the DRWR mechanism, given its statistical significance, leads to bias and is not advisable.

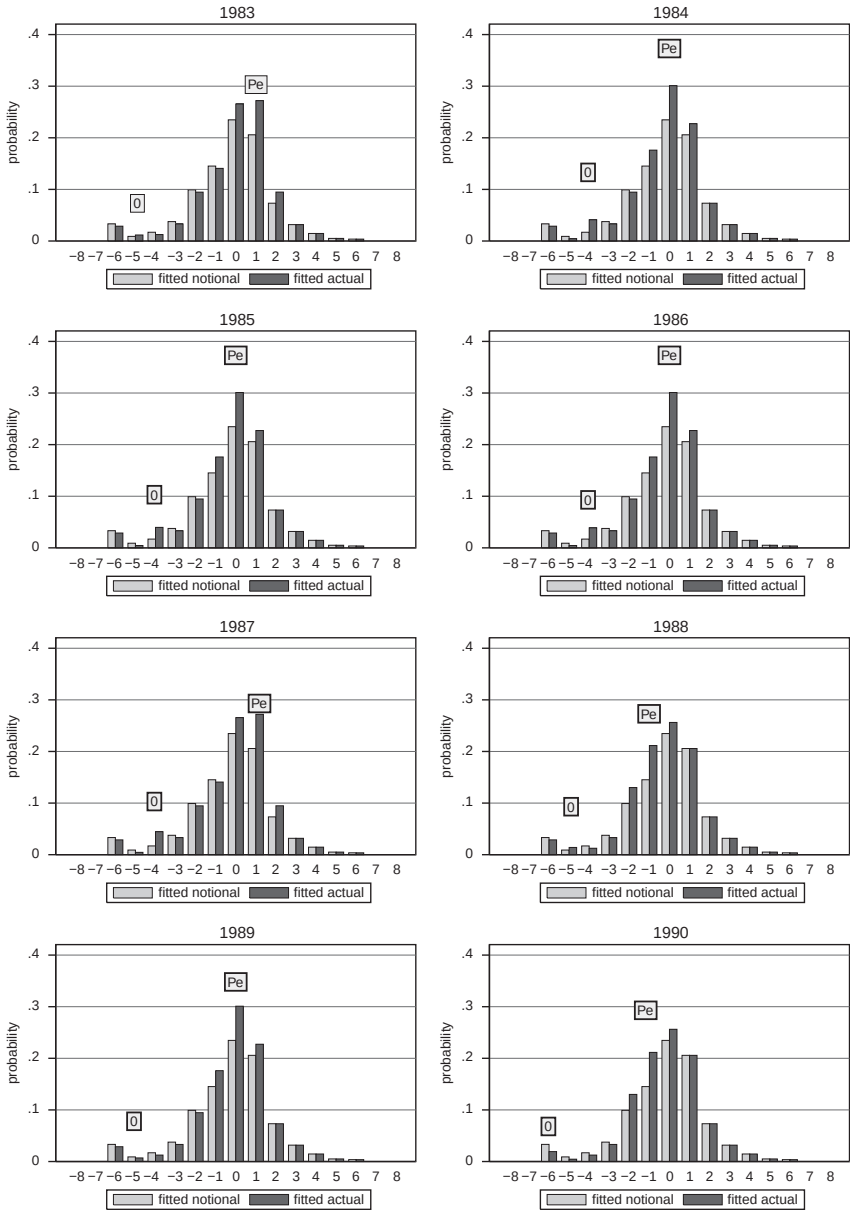

Fig. 7. Notional Versus Actual Nominal WGDs (Fitted Values): Medium-Inflation Period, FGLS Results in Table 5 (Both Types of Rigidity).

Table 8. Estimation Results: Low-Inflation Period.

Parameter	FGLS		Corrected OLS	
	Estimate	SE	Estimate	SE
β_{10}	0.1377*	0.0085	0.1490*	0.0094
β_{11}	0.0969*	0.0054	0.1143*	0.0090
β_{12}	0.0791*	0.0043	0.1110*	0.0074
β_{13}	0.0640*	0.0038	0.0544*	0.0046
β_{14}	0.0471*	0.0034	0.0444*	0.0043
β_{15}	0.0390*	0.0033	0.0400*	0.0027
β_{16}	0.0242*	0.0025	0.0218*	0.0022
β_{17}	0.0053*	0.0016	0.0041**	0.0017
β_{18}	0.0010	0.0014	−0.0058*	0.0015
β_{21}	−0.0191**	0.0076	−0.0219**	0.0105
β_{22}	0.0417*	0.0080	0.0171***	0.0100
β_{23}	0.0063	0.0062	0.0232*	0.0069
β_{24}	−0.0069	0.0052	−0.0008	0.0057
β_{25}	−0.0174*	0.0043	−0.0025	0.0045
β_{26}	−0.0125*	0.0032	0.0069***	0.0038
β_{27}	−0.0001	0.0020	0.0012	0.0022
β_{28}	0.0024	0.0017	0.0217*	0.0025
γ_1	0.3601*	0.0113	0.3159*	0.0165
γ_2	−0.1213*	0.0074	−0.1169*	0.0109
γ_{36}	−0.0199*	0.0022	−0.0201*	0.0021
γ_{35}	−0.0260*	0.0046	−0.0315*	0.0029
γ_{34}	−0.0455*	0.0034	−0.0475*	0.0036
γ_{33}	−0.0601*	0.0041	−0.0614*	0.0044
γ_{32}	−0.0748*	0.0043	−0.0860*	0.0059
γ_{31}	−0.0965*	0.0056	−0.1121*	0.0076
γ_4	0.0246*	0.0017	0.0294*	0.0021
δ_{-15}	0.0042	0.0029	0.0054**	0.0027
δ_{-14}	0.0102*	0.0034	0.0241*	0.0071
δ_{-13}	0.0040	0.0049	0.0067	0.0066
δ_{-12}	0.0207*	0.0054	0.0085	0.0072
δ_{-11}	0.0167**	0.0075	−0.0166***	0.0086
δ_{10}	0.0188*	0.0061	0.0027	0.0071
δ_{11}	0.0193*	0.0062	−0.0028	0.0070
N	102		102	

*Significance level at 1%.
**Significance level at 5%.
***Significance level at 10%.
Note: DNWR and DRWR effects are allowed for.

Table 9. Estimation Results: Low-Inflation Period.

Parameter	FGLS		Corrected OLS	
	Estimate	SE	Estimate	SE
β_{10}	0.1524*	0.0076	0.1449*	0.0086
β_{11}	0.1075*	0.0044	0.1141*	0.0070
β_{12}	0.0870*	0.0037	0.1081*	0.0069
β_{13}	0.0744*	0.0032	0.0625*	0.0041
β_{14}	0.0545*	0.0031	0.0457*	0.0042
β_{15}	0.0473*	0.0028	0.0411*	0.0025
β_{16}	0.0276*	0.0024	0.0219*	0.0021
β_{17}	0.0061*	0.0016	0.0037**	0.0017
β_{18}	0.0013	0.0014	−0.0061*	0.0015
β_{21}	−0.0185*	0.0071	−0.0186***	0.0097
β_{22}	0.0496*	0.0075	0.0178***	0.0100
β_{23}	0.0023	0.0060	0.0128***	0.0067
β_{24}	−0.0086***	0.0049	−0.0020	0.0058
β_{25}	−0.0245*	0.0039	−0.0041	0.0044
β_{26}	−0.0158*	0.0031	0.0068***	0.0038
β_{27}	−0.0009	0.0020	0.0016	0.0022
β_{28}	0.0022	0.0017	0.0220*	0.0025
γ_1	0.3838*	0.0100	0.3269*	0.0157
γ_2	−0.1375*	0.0067	−0.1147*	0.0107
γ_{36}	−0.0229*	0.0021	−0.0200*	0.0021
γ_{35}	−0.0328*	0.0043	−0.0321*	0.0027
γ_{34}	−0.0522*	0.0030	−0.0473*	0.0034
γ_{33}	−0.0683*	0.0036	−0.0606*	0.0039
γ_{32}	−0.0832*	0.0038	−0.0839*	0.0053
γ_{31}	−0.1006**	0.0043	−0.1014*	0.0057
γ_4	0.0273*	0.0015	0.0296*	0.0019
N	102		102	

*Significance level at 1%.
**Significance level at 5%.
***Significance level at 10%.
Note: DRWR effects are suppressed.

Fig. 8 plots the estimated (FGLS) probability histograms for the notional and actual WGDs based on Table 8.

6. CONCLUSION

In this chapter, we explored several improvements to the method of constructing the stage 1 histograms that underly the estimation, in stage 2,

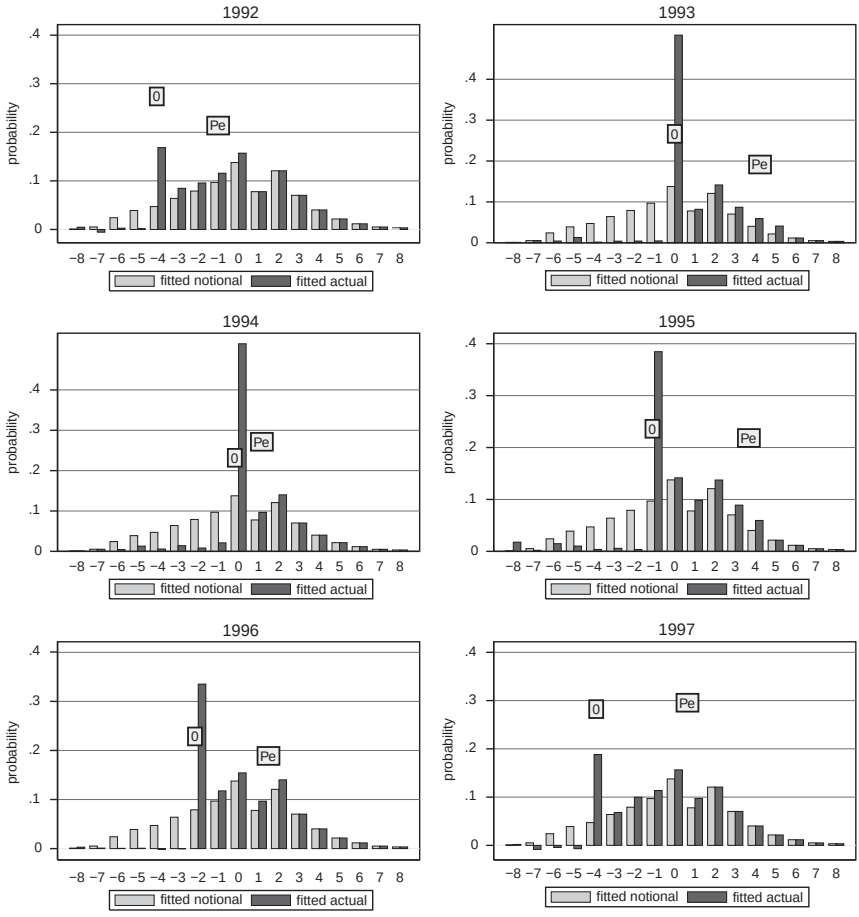

Fig. 8. Notional Versus Actual Nominal WGDs (Fitted Values): Low-Inflation Period, FGLS Results in Table 8 (Both Types of Rigidity).

of DNWR and DRWR distortions. The conceptual improvements (centering on the median of the WGD and using Kernel methods) produce stage 1 data that are very similar to data from the relative frequency approach. In the stage 2 sub-period estimations, the model performed as expected, failing to find DNWR in the high-inflation period but confirming the existence of distortions due to DNWR and DRWR in the medium- and low-inflation periods. An interesting issue is whether the DRWR mechanism, present in the medium- and low-inflation periods, would suggest that its

omission (as in earlier studies) would qualify the results obtained for DNWR. Suppressing DRWR does, indeed, modify the estimates for the spike at zero under DNWR. It is underestimated in the medium and over-estimated in the high-inflation period, confirming that omitting important variables is not advisable. Of course, other estimates of the DNWR mechanism (e.g. how the spike at zero diminishes as the median of the WGD increases) are also biased when the importance of DRWR is suppressed.

A particular challenge has been the identification of DRWR during the high-inflation period. Our method, being data-driven and, essentially, semi-parametric, relies on there being sufficient differentiation in the relation between the median of the WGD and the mean of the AID. This may be one reason why our estimates are not well-identified. A related point is that the mass that, due to DRWR, is shifted towards the centre of the AID is larger if the expected inflation rate is high relative to the centre of gravity of the WGD. Table 1 shows that this is only true of one of the years in the 1977–1982 period, thus limiting the quantitative significance of DRWR. Finally, this being the period of adverse oil price shocks, may explain why more moderate wage growth would have been acceptable. Clearly more research in these important issues is warranted.

NOTES

1. In the context of more contemporary models, where productivity shocks shift the labour demand curve, the real wage rate may be procyclical.
2. As an extreme example, in the case of firm and uniform inflation expectations, where absolutely all agents are subject to DRWR (interpreted to mean that no one will accept a real wage cut), the issue of DNWR becomes moot – except when deflation is expected. Only then will the DNWR mechanism be relevant at values of wage adjustment that exceed the expectation of inflation. Under less stringent conditions, for example when the anticipated inflation distribution (AID) is not degenerate and contains the point zero, it may be necessary to specify whether DNWR or DRWR takes precedence. Suppose, for instance that an agent expects inflation to be -1%, is offered a -3% wage adjustment (i.e. is subject to both a nominal and a real cut); in such a case, will the line of resistance be drawn at zero nominal adjustment (DNWR and an implied anticipated real wage increase of 1%), or at -1% nominal adjustment (DRWR and an implied real wage constancy)? Depending on how the question of which mechanism takes precedence is resolved, this will be reflected in the actual wage adjustment outcomes and the ability to distinguish the processes involved.
3. See, for example Silverman (1986) and Wasserman (2006) for discussion.
4. See Li and Racine (2007).

5. The latter decision is the most critical. The estimation was carried out in *R*, using the '*np*' package. We are grateful to Qi Li for information and to Jeff Racine for code that implements these procedures.

6. Because of the small number of contracts involved, the first two and the last three years in the sample are considered together in everything that follows and we refer to these as 'years' 1977 and 1997, respectively.

7. This is the one-year-ahead forecast from an AR(6) regression model with a GARCH(1,1) error process. This process also supplies the variance of the anticipated inflation rate at each point in time.

8. These are median centered, as discussed in Section 4.1.

9. Now $t = 1, \ldots, T$ becomes the observation index.

10. Both estimators satisfy this requirement asymptotically, and the relative frequency estimator also in finite samples.

11. In such a case, the system would consist of $2J + 1$ equations. The dependent variable corresponding to the equation for a particular observation would be $\hat{P}_{jt}$, where j is the equation index, and t the within equation observation index. To estimate the system we would have in total $(2J + 1) \times T$ observations, with T observations on each equation.

12. Given the parameterisation of the probability histograms under the alternative discussed below, we take the notional distribution to be the nominal WGD free of any DWR or menu-cost distortions.

13. This would imply a positive relationship between the spread and location of the histograms of the actual-wage-growth data irrespective of whether DWR is present or not.

14. The assumption in the original Kahn (1997) methodology that the shape of the notional distribution is the same across years, has often been cited as one of the main drawbacks of this methodology as in most actual-wage-growth datasets there appears to exist a variation in the spread of the distribution across years characterised by different levels of inflation. This point is raised by Nickell and Quintini (2003) who go on to propose a flexible way of studying DNWR.

15. The index k is assumed to take values from the set $\{k_{\min}, \ldots, 0, \ldots, k_{\max}\}$. The bin for which $k = 0$ contains the centre of the AID, bins with positive values of k are located to the right of this bin, and bins with negative values to its left. The values taken by $k_{\min}$ and $k_{\max}$ are determined empirically.

16. The sub-periods are defined with respect to the values of the estimated mean anticipated inflation, as it is the AID that determines the nature of distortions due to DRWR.

17. The corresponding coefficients are γ_{31} to γ_{36}.

ACKNOWLEDGMENT

We thank M. Legault, Human Resources Development Canada, for the data and the Social Sciences and Humanities Research Council for financial support.

REFERENCES

Abraham, K. G., & Haltiwanger, J. C. (1995). Real wages and the business cycle. *Journal of Economic Literature*, *XXXIII*, 1215–1264.

Altonji, J. G., & Devereux, P. J. (2000). The extent and consequences of downward nominal wage rigidity. In: S. W. Polachek (Ed.), *Research in labor economics* (Vol. 19, Chapter 10, pp. 383–431). Greenwich, CT: JAI Press Inc.

Barwell, R., & Schweitzer, M. E. (2007). The incidence of nominal and real wage rigidities in Great Britain: 1978–98. *Economic Journal*, *117*(524), F553–F569.

Bauer, T. K., Bonin, H., Goette, L. F., & Sunde, U. (2007). Real and nominal wage rigidities and the rate of inflation: Evidence from West German micro data. *Economic Journal*, *117*(524), F508–F529.

Bewley, T. F. (1999). *Why wages do not fall during a recession?* Cambridge: Harvard University Press.

Card, D., & Hyslop, D. (1997). Does inflation grease the wheels of the labor market? In: C. Romer & D. Romer (Eds), *Reducing inflation: Motivation and strategy* (pp. 114–121). Chicago: University of Chicago Press.

Christofides, L. N., & Leung, M. T. (2003). Nominal wage rigidity in contract data: A parametric approach. *Economica*, *70*(280), 619–638.

Christofides, L. N., & Li, D. (2005). Nominal and real wage rigidity in a friction model. *Economics Letters*, *87*, 235–241.

Christofides, L. N., & Nearchou, P. (2007). Real and nominal wage rigidities in collective bargaining agreements. *Labour Economics*, *14*, 695–715.

Christofides, L. N., & Stengos, T. (2001). A non-parametric test of the symmetry of PSID wage-change distributions. *Economics Letters*, *71*, 363–368.

Christofides, L. N., & Stengos, T. (2002). The symmetry of the wage-change distribution: Survey and contract data. *Empirical Economics*, *4*, 705–723.

Christofides, L. N., & Stengos, T. (2003). Wage rigidity in Canadian collective bargaining agreements. *Industrial and Labor Relations Review*, *56*(3), 429–448.

Cramton, P., & Tracy, J. S. (1992). Strikes and holdouts in wage bargaining: Theory and data. *American Economic Review*, *82*, 100–121.

Crawford, A., & Harrison, A. (1998). Testing for downward rigidity in nominal wage rates. In: *Price stability inflation targets and monetary policy* (pp. 179–225). Ottawa: Bank of Canada.

Dickens, W., & Groshen, E. (2004). The International Wage Flexibility Project (IWFP). Proceedings of the final conference, European Central Bank, Frankfurt Am Main, Germany.

Dunlop, J. T. (1938). The movement of real and money wages. *Economic Journal*, *48*, 413–434.

Fortin, P. (1996). The great Canadian slump. *Canadian Journal of Economics*, *29*(4), 761–787.

Holden, S. (1989). Wage drift and bargaining: Evidence from Norway. *Economica*, *56*(224), 419–432.

Holden, S. (1994). Wage bargaining and nominal rigidities. *European Economic Review*, *38*, 1021–1039.

Holden, S. (1998). Wage drift and the relevance of centralised wage setting. *Scandinavian Journal of Economics*, *100*, 711–731.

Holden, S. (2004). The costs of price stability: Downward nominal wage rigidity in Europe. *Economica*, *71*, 183–208.

Holden, S., & Wulfsberg, F. (2007). Downward nominal wage rigidity in the OECD. Working Paper Series 777. European Central Bank, Frankfurt.

Kahn, S. (1997). Evidence of nominal wage stickiness from microdata. *American Economic Review*, *87*(5), 993–1008.

Lebow, D. E., Stockton, D. J., & Wascher, W. L. (1995). *Inflation, nominal wage rigidity, and the efficiency of labor markets*. Finance and Economics Discussion Series 1995–45. Washington, DC: Board of Governors of the Federal Reserve System.

Li, Q., & Racine, J. (2007). *Nonparametric econometrics*. Princeton: Princeton University Press.

Macleod, W. B., & Malcomson, J. M. (1993). Investment, holdup, and the form of market contracts. *American Economic Review*, *37*, 343–354.

Malcomson, J. M. (1997). Contracts, hold-up, and labor market. *Journal of Economic Literature*, *35*(4), 1916–1957.

McLaughlin, K. J. (1994). Rigid wages? *Journal of Monetary Economics*, *34*, 383–414.

Nickell, S., & Quintini, G. (2003). Nominal wage rigidity and the rate of inflation. *The Economic Journal*, *113*, 762–781.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. New York, NY: Chapman and Hall.

Smith, J. (2000). Nominal wage rigidity in the United Kingdom. *The Economic Journal*, *110*, C176–C195.

Solon, G., Barsky, R., & Parker, J. (1994). Measuring the cyclicality of real wages: How important is composition bias. *Quarterly Journal of Economics*, *109*(1), 1–25.

Tarshis, L. (1939). Changes in real and money wages. *Economic Journal*, *49*, 150–154.

Wasserman, L. (2006). *All of nonparametric statistics*. New York, NY: Springer.