Essays on the Task Content of Occupations and Occupational Mobility

by

Guido Matías Cortés

B.Econ., Universidad de Costa Rica, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Economics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

July 2012

Abstract

This dissertation studies the effects of technological change on workers' occupational choices and wages, as well as the human capital costs associated with occupational transitions. The first part of the dissertation focuses on the interaction between technological change and tasks. Over the past three decades technological improvements have led to a dramatic reduction in the employment share of occupations with a high content of routine tasks in the United States and other developed countries. This dissertation provides a novel perspective on this phenomenon by focusing on the individual-level effects of this type of technological change in terms of occupational switching patterns and wage changes. I formalize the predicted effects within the context of a model of occupational sorting based on comparative advantage, and I test them using data from the Panel Study of Income Dynamics (PSID) from 1976 to 2007. Consistent with the predictions of the model, I find strong evidence of selection on ability in the occupational mobility patterns of workers in routine occupations, with those of relatively high (low) ability switching to non-routine cognitive (non-routine manual) occupations. In terms of wage growth, also consistent with the prediction of the model, workers in routine jobs experience significant declines in their wage premia relative to workers in any other type of occupation. Switchers from routine to either type of non-routine job (cognitive or manual) experience significantly higher wage growth than stayers over long-run horizons. The second portion of the dissertation analyzes the role of the task content of occupations. I develop a measure of task distance between occupation pairs and study its impacts from two different perspectives: At a microeconomic level, I analyze the wage changes for workers experiencing occupational transitions of different distances. At a macroeconomic level, I analyze the impacts of task distance on the aggregate flows of workers across occupations. The aggregate-level evidence suggests that the cost of switching occupations is increasing in distance, but only for switches occurring across broad occupation groups. The individual-level evidence suggests that there is a negative correlation between wage changes and distance, but only for certain subsets of workers.

Table of Contents

Abstract ii

Table of Contents iii

List of Tables vi

List of Figures viii

Acknowledgments x

1 Introduction 1

2 Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data 8

2.1 Introduction 8
2.2 Model 14
2.2.1 Household Preferences 15
2.2.2 Firms 15
2.2.3 Labor Productivity 16
2.2.4 Worker Sorting and Wages 16
2.2.5 Equilibrium 17
2.3 Effects of Routine-Biased Technical Change 18
2.3.1 Switching Patterns Induced by RBTC 19
2.3.2 Wage Changes Induced by RBTC 21
2.3.3 Summary of the Predictions of the Model 23
2.4 Data 25
2.5 Empirical Implementation 29
2.6 Results: Effects of Routine-Biased Technical Change 33
2.6.1 Switching Patterns 33
2.6.2 Wage Changes 41
2.7 Robustness Checks 48
2.7.1 Time-Varying Skills 48

2.7.2 Changing Returns to Education 50
2.7.3 Alternative Classification of Occupations 55
2.8 Conclusions 59

3 Task Distance: Construction, Trends and Relationship with Wage Changes for Individual Workers 61
3.1 Introduction 61
3.2 Data and Methodology 64
3.2.1 PSID Data 65
3.2.2 DOT Data 66
3.2.3 Matching DOT with PSID 67
3.2.4 Selection of Occupational Characteristics 68
3.2.5 Distance Measures 68
3.3 Findings: Trends and Cross-Sectional Variation in Task Distance 69
3.4 Relationship Between Task Distance and the Wage Changes of Individual Workers 74
3.5 Conclusions and Future Work 79

4 The Barriers to Occupational Mobility: An Aggregate Analysis 80
4.1 Introduction 80
4.2 Model 83
4.2.1 Setup 83
4.2.2 Wages and Productivity Draws 84
4.2.3 Flows of Workers Across Occupations 85
4.3 Empirical Implementation 87
4.4 Data 89
4.5 Results 90
4.5.1 Gravity Equation Estimation 90
4.5.2 Estimation of θ 92
4.5.3 Implied Mobility Costs 97
4.5.4 Counterfactual Experiments 99
4.6 Conclusion 101

5 Conclusion 102

Bibliography 104

Appendices

A Appendices to Chapter 2 110
A.1 Proposition Proofs 110
A.1.1 Proof of Proposition 1 110
A.1.2 Proof of Proposition 2 111
A.1.3 Proof of Proposition 3 112
A.2 Model Extension: Two-Dimensional Skills 113
A.3 Grouping of Occupation Codes 115

B Appendices to Chapter 3 118
B.1 Robustness Check - Evolution of Distance with O*Net Measures 118
B.1.1 O*Net Data 118
B.1.2 Relevant Occupational Characteristics 119
B.1.3 Matching O*Net Data with the PSID 121
B.1.4 Adjustment for Coding Error 121
B.1.5 Findings 122
B.1.6 Patterns for Different Age and Education Groups . . 127
B.2 Analysis of CPS Data 128
B.2.1 CPS Data 129
B.2.2 Matching DOT/O*Net with CPS 132
B.2.3 Results 134
B.3 Description of PSID Variable Construction 138

C Appendices to Chapter 4 140
C.1 Matching DOT with CPS 141
C.2 Extension: Occupation Tenure 143

List of Tables

2.1 Descriptive statistics (1976-2007) 27
2.2 Regressions of the probability of switching occupations two years ahead on dummies for ability quintiles (routine workers, odd years only) 36
2.3 Regressions of the probability of switching to particular occupations two years ahead for routine workers (odd years only, 1977-2005) 38
2.4 Regressions of the probability of switching to particular occupations two years ahead for routine workers (odd years only) 39
2.5 Regression of changes in log real wages over different time horizons on dummies for initial occupation 42
2.6 Regression of changes in log real wages over different time horizons on dummies for initial occupation for workers who do not switch occupations 43
2.7 Wage changes for routine workers, according to direction of switch 47
2.8 Robustness checks for the wage changes for routine workers, according to direction of switch 49
2.9 Regressions of the probability of switching two years ahead for routine workers (odd years only, 1977-2005) 54
2.10 Wage changes for routine workers, according to direction of switch 56

3.1 Examples of DOT-77 scores (rescaled to 0-10 range) 68
3.2 Examples of some observed low- and high-distance occupational transitions 69
3.3 Regression of distance on observables - PSID data 73
3.4 Autocorrelation of wages and relation to distance - Dependent variable: Log current hourly real wage 75
3.5 Estimated auto-correlation of wages at different percentiles of distance 75

3.6 Estimated marginal effect of distance on current wages at different percentiles of lagged wages 76
3.7 Wage change regressions for occupational switchers 77
3.8 Wage change regressions 78

4.1 Examples of DOT-91 scores 91
4.2 Estimated coefficients on 'gravity-type' equation, 1983-2002 . 92
4.3 Summary statistics of estimates of θ 97
4.4 Estimated effects on occupational transition costs for each of the specifications in Table 4.2 98
4.5 Estimated barriers to occupational transition for some commonly observed transitions (OLS specification excluding occupation pairs with zero flows) 99
4.6 Ratio of counterfactual mobility to fitted mobility for selected occupations 100

A.1 Occupation code groupings 116

B.1 Conversion of DOT data to different coding levels 134
B.2 Conversion of O*Net data to different coding levels 136

C.1 2-digit occupation groupings for the Meyer and Osborne (2005) coding system 140
C.1 2-digit occupation groupings for the Meyer and Osborne (2005) coding system 141
C.2 Estimated coefficients on 'gravity-type' equation including mean tenure as a control, 1983-2002 148

List of Figures

2.1 Effects of routine-biased technical change on skill cutoffs and wages (case where C_{cog} increases relative to C_{man}) 24
2.2 Changes in employment shares for broad occupation groups, men, PSID, 1968-2007 28
2.3 Exit probabilities by ability quintile 35
2.4 Direction of switch by ability quintile 37
2.5 Direction of switch by ability quintile, 1976-1979 40
2.6 Estimated coefficients on occupation-year dummies 44
2.7 Estimated coefficients on occupation-year dummies when controlling for time-varying skills (age bins) 51
2.8 Education effects 53
2.9 Estimated coefficients on occupation-year dummies 55
2.10 Direction of switch by ability quintile 57
2.11 Estimated coefficients on occupation-year dummies under alternative classification of occupations (1976-2001) 58

3.1 Median distance for occupational switchers, all household heads and wives 71

4.1 Scatter plot of estimated source and destination occupation fixed effects, OLS specification **excluding** occupation pairs with zero flows 93
4.2 Scatter plot of estimated source and destination occupation fixed effects, OLS specification **including** occupation pairs with zero flows 94
4.3 Scatter plot of estimated source occupation fixed effects from the OLS specification **including** occupation pairs with zero flows against log occupation size 95
4.4 Counterfactual mobility with no cost of transition 100

B.1 Mean annual distance for occupational switchers adjusted for the change in the coding procedure 123

B.2 Mean annual distance adjusted for coding error for alternative measures of distance . . . 124
B.3 Mean annual distance adjusted for coding error - Distance measures based on importance scores of different groups of abilities as indicated in each graph . . . 125
B.4 Mean annual distance adjusted for coding error - Distance measures based on importance scores of different groups of work activities as indicated in each graph . . . 126
B.5 Adjusted mean distance of occupational switches (conditional on switching) by age and education groups . . . 127
B.6 Mean distance of occupational switches (conditional on switching) by age for different cohorts and education groups . . . 128
B.7 Periods with consistent data on occupational transitions from the CPS and the PSID . . . 131
B.8 Graphic summary of procedure followed to aggregate DOT and O*Net codes to the different COC-code levels . . . 135
B.9 Median distance for occupational switchers, CPS, all household heads . . . 136
B.10 Median distance for occupational switchers, CPS, only male private-sector household heads . . . 137

C.1 Histogram of estimated values of θ . . . 142

Acknowledgments

I am grateful to many people who in different ways supported me throughout the PhD program and without whom I would not have been able to complete this dissertation. First, I would like to thank my co-supervisors Nicole Fortin and Giovanni Gallipoli for always being available to answer my questions and concerns, for their patience, and for their exceptional guidance throughout the research process and in successfully navigating the job market. I also thank David Green, Thomas Lemieux, Henry Siu, Keith Head, and other faculty members and fellow students at UBC for very helpful discussions and feedback on the research in this dissertation. I also thank my professors at the Universidad de Costa Rica, particularly Jose Antonio Cordero and Anabelle Ulate, for encouraging me to pursue graduate studies, and for helping me arrive in Vancouver well prepared to tackle the first-year PhD courses. I would like to express my gratitude to the Canadian Labour Market and Skills Researcher Network (CLSRN) for their continued financial support throughout my PhD, as well as to Maureen Chin, Kathy Meredith and Maria Smith for always diligently taking care of all of the administrative aspects of the PhD and the job market. Finally, and perhaps most importantly, I am extremely grateful to my family, my partner, and my good friends for their unconditional support and their continuous encouragement, particularly during difficult times. Having so many great people in my life helped keep me balanced, happy and ultimately more productive.

Chapter 1

Introduction

This dissertation studies the effects of technological change on workers' occupational choices and wages, as well as the human capital costs associated with occupational transitions, with a specific focus on the task content of occupations. The first research chapter in the dissertation, Chapter 2, focuses on the role of task-replacing technologies. It studies what has happened to individuals working in occupations that have largely disappeared due to the arrival of new technologies. In the remaining two research chapters of the dissertation the focus is shifted towards the costs associated with job transitions involving different amounts of task switching. These chapters develop a measure of task distance and attach it to observed occupational transitions in the data. Chapter 3 performs a microeconomic level analysis, focusing on the wage changes for workers experiencing occupational transitions of different distances. Chapter 4 provides a macroeconomic level analysis, focusing on the impacts of task distance on the aggregate flows of workers across occupations. Each Chapter in the dissertation is written as a separate paper and may be read independently. The theme linking all the Chapters is the focus on the task content of occupations and its implications for human capital and wages.

Recent literature has linked technological change and changes in the wage distribution to the task content of occupations. Since the late 1980s, the labor market in the United States and other developed countries has become increasingly polarized. The share of employment in high-skill, high-wage occupations and in low-skill, low-wage occupations has been increasing relative to the share in occupations in the middle of the distribution. At the same time, wages have grown faster at the top and the bottom of the distribution than in the middle sections (Acemoglu and Autor, 2011). Pioneering work by Autor, Levy, and Murnane (2003), Autor, Katz, and Kearney (2006) and Goos and Manning (2007) has linked the polarization phenomenon to the occupational structure of the economy, and in particular to the task content of different occupations. Workers in the middle of the wage distribution tend to be concentrated in occupations with a high content of routine tasks, as measured by information in the Dictionary of Occupational Titles (DOT).

At the same time, technological changes occurring since the 1980s have resulted in the creation of capital, such as machines and computers, that can perform mainly routine tasks and can therefore substitute for workers in occupations with high routine task content. This hypothesis has become known as 'routinization', or routine-biased technical change (RBTC).

In this dissertation, I investigate the implications of routine-biased technical change within the context of a general equilibrium model with endogenous sorting of workers into occupations based on comparative advantage. The novel aspect in my research is the focus on the individual-level predictions in terms of occupational switching patterns and wage changes. The approach taken in this dissertation provides micro-level evidence on the dynamics underlying the aggregate patterns of employment and wage polarization, on the way in which particular subsets of workers have been impacted by routinization, and on the changes over time in occupational wage premia once selection has been accounted for. To the best of my knowledge, this is the first paper to directly use individual-level panel data to study the labor market experience of routine workers in the U.S. over the past three decades, thus shedding light on what has happened to these workers over time.[1]

The occupational sorting mechanism featured in the model follows Gibbons, Katz, Lemieux, and Parent (2005): workers select into occupations based on their comparative advantage. Capital is modeled as suggested by Autor, Levy, and Murnane (2003): it enters the production function as a substitute for labor working in routine tasks, and a complement for workers in non-routine cognitive tasks. I derive the model's predictions for the effects of routine-biased technical change (RBTC). The model makes the following predictions: RBTC induces workers at the bottom of the ability distribution within routine occupations to switch to non-routine manual jobs, while inducing those at the top end to switch to non-routine cognitive jobs. The model also makes predictions in terms of the changes in occupational wage premia: The wage premium in routine occupations is predicted to fall relative to that in the two non-routine occupational categories. For this reason, workers staying in routine jobs experience a fall in wages, relative to those staying in either non-routine manual or non-routine cognitive jobs. Within the context of the model, switchers must do at least as well as stayers in terms of wage growth.

To test the predictions of the model for individual workers, the paper

[1]A common assumption has been that most of these workers have been displaced into low-skill jobs (e.g. Acemoglu and Autor (2011), p.64), but there has been little evidence put forth to support this claim.

uses data from the Panel Study of Income Dynamics (PSID). Occupations are grouped based on an aggregation of 3-digit occupation codes into the three categories used in the model: non-routine manual (service occupations), routine (sales, clerical, craftsmen, foremen, operatives, laborers), and non-routine cognitive (professional and managerial occupations). The empirical strategy involves the estimation of a wage equation derived from the model. An individual worker's potential wage in each occupation consists of an occupation-specific premium (common to all workers in the same occupation in a given year), as well as an occupation-specific return to the worker's skills. Skills are allowed to contain both observable and unobservable components. Workers select into the occupation where their potential wage is highest. The key identifying assumptions for the estimation of the wage equation are that unobservable skills are time-invariant, workers have full information about their skills, and any idiosyncratic temporary shocks to individual wages are independent of sectoral choice. Under these assumptions, estimating a wage equation with *occupation spell* fixed effects (interactions of individual fixed effects with occupation dummies) controls for the self-selection of workers into occupations based on unobserved ability, and allows for consistent estimation of the changes over time in occupation wage premia. The estimated occupation spell fixed effects themselves are also informative, as they can be used to rank workers according to ability within occupation-year cells.[2]

Using this empirical strategy, I find that there is strong evidence of selection on ability for workers switching out of routine jobs: Low ability routine workers are more likely to switch to non-routine manual jobs, while high ability routine workers are more likely to switch to non-routine cognitive jobs. This is fully consistent with the predictions of the model.

In terms of wage growth, I find that workers staying in routine jobs perform significantly worse than workers staying in any other type of occupation. The wage premium for routine occupations is estimated to have fallen by 17% from 1976 to the mid-2000s, relative to the wage premium for non-routine manual occupations (14% when taking account of changing returns to education). Meanwhile, over the same time period, the wage premium for non-routine cognitive occupations is estimated to have risen by 25% (7% when taking account of changing returns to education) relative to the wage premium for non-routine manual occupations.

There are also significant differences in wage growth between routine

[2]The empirical strategy may be extended to allow for changes over time in the return to education and the empirical results are robust to this extension. See Section 2.7.2.

workers who stay or switch to other occupations: Those who switch to non-routine manual jobs have significantly lower wage growth than stayers over short run horizons (around 14% lower over a two-year period), but subsequently recover from these losses and have significantly faster wage growth than stayers in the long run (5 to 12% higher over a 10-year period). Meanwhile, those who switch to non-routine cognitive jobs have significantly higher wage growth than stayers over a variety of time horizons (6 to 12% higher over a two-year period; 14 to 16% higher over a 10-year period).

In the remainder of the dissertation the focus is shifted towards the human capital costs associated with job transitions involving different amounts of task switching. Understanding how individuals build their human capital and how transferable this human capital is across jobs has been an issue of great interest in economics. The extent to which an individual's human capital is specific to a firm or an industry, for example, or general (and thus transferable across jobs) has important implications for the analysis of an economy's reaction to shocks, and for the definition of appropriate policy responses to issues such as unemployment.

Several contributions to the human capital literature have analyzed the costs associated with different types of employment transitions. Topel (1991) provides evidence that an important component of wages of male workers in the United States is related to firm tenure. Dustmann and Meghir (2005), using data for West Germany, also find substantial returns to firm tenure, particularly for unskilled workers. Meanwhile, Neal (1995) and Parent (2000), suggest that workers lose a large fraction of their human capital when they switch industries, rather than firms. More recently, Kambourov and Manovskii (2009b) have shifted the focus towards occupation-specific human capital. They argue that firm and industry tenure play only minor roles in determining wages once occupational tenure has been accounted for. Support for this argument, at least among certain occupational groups, is also found in Sullivan (2010).

The findings in Kambourov and Manovskii (2009b), along with the evidence they present in Kambourov and Manovskii (2008), suggesting that occupational mobility in the U.S. is high and rose significantly between 1968 and 1997, motivate a deeper study of the human capital costs associated with occupational mobility. In this dissertation, I argue that an adequate analysis of the costs of occupational mobility must take into account that occupational switches (which empirically are based on whether the occupation code assigned to a worker changes) encompass many different types of transitions. In some cases a worker may be completely changing careers, while in others it may involve only a minor change in the set of tasks performed by the worker. Lazear (2003), Poletaev and Robinson (2008) and Gathmann and Schonberg (2010) argue in favor of linking human capital to tasks, and therefore differentiating occupational transitions according to the extent of task switching that they entail. Human capital built in an occupation should be partially transferable to other occupations where the set of tasks performed is similar.[3]

In this dissertation, I characterize occupations by a set of skill requirements using information from the Dictionary of Occupational Titles (DOT), and follow Gathmann and Schonberg (2010) in developing a distance measure between occupations. With this distance measure, occupational mobility can be characterized in more detail, by distinguishing between moves among occupations with different degrees of similarity in their skill requirements. Chapter 3 of the dissertation analyzes the evolution over time in the distance of observed occupational transitions using data from the Panel Study of Income Dynamics (PSID) from the United States for the period 1968-1997. In an environment of rising occupational mobility, I address the question of whether the increase in occupational mobility documented in Kambourov and Manovskii (2008) has been driven by transitions between more similar or more different jobs. The chapter also analyzes the impact of the distance of occupational switches on wages. If human capital is specific to a set of skills, one would expect workers to experience larger wage losses if they experience a larger change in the mix of tasks they perform.[4]

I find evidence of a significant but very small increase in the distance of occupational switches between 1968 and 1997. I do not find a robust negative correlation between the distance switched and the wage changes of individual workers. Instead, I find some evidence that wage changes are negatively correlated with distance only for particular subsets of workers,

[3]Lazear (2003) suggests that there exist a set of skills that are general, in the sense that they are used in many firms. Each firm, however, uses a different combination of these skills. What is firm-specific is then the particular weighting of different skills. Gathmann and Schonberg (2010) develop the concept of "task-specific human capital": skills accumulated in an occupation are only productive in occupations in which a similar combination of tasks is performed. They use data from Germany to generate distance measures based on the (dis)similarity of the mix of tasks performed in different occupations, and find that task-specific human capital accounts for an important fraction of wage growth. Meanwhile, Poletaev and Robinson (2008) follow Ingram and Neumann (2006) in estimating a factor model to determine the "basic skills" that define an occupation. Their data source is the Dictionary of Occupational Titles (DOT) of the United States. They define a discrete measure of skill switching and apply it to the set of occupational switchers observed in the Displaced Worker Survey of the U.S. The authors find that the wage cut for 'skill switchers' is significantly larger than for 'skill stayers'.

[4]See Robinson (2011) for a recent paper using a similar approach.

namely workers who experience occupational downgrades, defined as decreases in their intensity of abstract non-routine task usage.

Chapter 4 of the dissertation aims to provide an estimate of the costs of occupational mobility by using data on the aggregate flows of workers across occupations. The approach taken in the Chapter is somewhat unconventional relative to previous literature. I adapt a framework which is widely used in the trade literature in the form of 'gravity models'. In that literature, the interest is in estimating barriers to trade using data on flows of goods across countries, and proxies for trade costs that include geographical distance and whether the countries share a common border or a common language, among others. I show how the framework may be reinterpreted in order to consider the barriers to occupational mobility using data on the flows of workers across occupations, and proxies for mobility costs, mainly based on task data. Within the context of the model, the barriers to occupational mobility involve losses of human capital that are incurred when a worker needs to adapt the set of tasks that he is familiar with, or learn a new set of tasks altogether, upon switching occupations.

Previous models of occupational choice, such as Gibbons, Katz, Lemieux, and Parent (2005) or Groes, Kircher, and Manovskii (2009), have focused on the roles of comparative advantage across skill groups and of learning in driving selection into occupations. In Kambourov and Manovskii (2009a) occupational mobility is driven by persistent shocks to occupations that impact productivity of all workers in the occupation equally. The model in this dissertation differs from the previous literature in that the driving force behind occupational mobility is a set of non-persistent individual-specific productivity draws.

The framework in the paper involves a static, partial equilibrium model with perfect information. There is a continuum of homogeneous workers, who only differ in terms of the occupation that they start the period in (which is exogenously pre-determined). Workers make productivity draws from a set of extreme value distributions for each potential occupation. Once workers observe their draws, they decide which occupation to work in during the period. There exist costs to switching occupations, which depend on the particular occupation that the worker starts in, and the particular occupation that he considers switching to. This switching cost is not individual-specific, but rather depends only on the identity of the worker's current and potential occupation. Based on his productivity draws, and taking into account the costs of mobility, the worker chooses the occupation where he will receive the highest wage. The workers' optimal switching decisions and the properties of the extreme value distribution lead to a 'gravity-type' equation, which predicts how the flow of workers across any given occupation pair is related to a set of occupation-specific characteristics, and to the cost of switching.

In order to estimate this gravity equation empirically, and to obtain an estimate of the barriers to occupational mobility, I need to define proxies for the costs of occupational mobility. As the main proxy, I use the measure of task distance developed and discussed in Chapter 3. I also consider the role of two other proxies that might reflect task similarities not captured by the distance measure. The first is whether the two occupations share the same major occupational category, and the second is whether they share the same type of main task.

I then estimate the gravity equation empirically using data on worker flows across occupations from the Current Population Survey (CPS) from 1983 to 2002. Task distance is found to be a very important determinant of the cost of switching occupations, suggesting an important role for task-specific human capital. A conservative specification of the implied effects of task distance on the cost of switching suggests that a one standard deviation of distance increases the cost of switching occupations by 16.3% if the occupations are in different major occupational groups and do not share the same type of main task. The effect is reduced to 15% if the occupation pair share the same type of main task. If the occupations are in the same major occupation group, however, the implied cost of the transition is negligible under this specification. The estimates suggest that aggregate occupational mobility rates would be approximately 1.5 times higher if there were no costs to switching occupations.

Chapter 2

Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data[5]

2.1 Introduction

Since the late 1980s, the labor market in the United States and other developed countries has become increasingly polarized. The share of employment in high-skill, high-wage occupations and in low-skill, low-wage occupations has been increasing relative to the share in occupations in the middle of the distribution. At the same time, wages have grown faster at the top and the bottom of the distribution than in the middle sections (Acemoglu and Autor, 2011).[6] Pioneering work by Autor, Levy, and Murnane (2003), Autor, Katz, and Kearney (2006) and Goos and Manning (2007) has linked the polarization phenomenon to the occupational structure of the economy, and in particular to the task content of different occupations. Workers in the middle of the wage distribution tend to be concentrated in occupations with a high content of routine tasks, as measured by information in the

[5]This chapter benefited from discussions with Nicole Fortin, Giovanni Gallipoli, David Green and Thomas Lemieux, as well as with participants at the UBC Empirical Lunch, the CEA Conference (2011), the SOLE Conference (2012), and seminar participants at SFU, Bank of Canada, Université de Montréal, UCL, Carleton, SOFI, University of Manchester, University of Queensland, University of Adelaide and Banco Central de Costa Rica. Financial support from CLSRN is gratefully acknowledged.

[6]The polarization phenomenon has also been documented for European countries; see Dustmann, Ludsteck, and Schönberg (2009) and Goos, Manning, and Salomons (2009). See also Gregory and Vella (1995) for earlier work on the disappearance of middle-wage occupations in Australia over the period 1976-1990. Employment share polarization has also been documented at a sub-national level for the state of California in Milkman and Dwyer (2002). Their work also highlights differences in the patterns observed across metropolitan areas within the state.

Dictionary of Occupational Titles (DOT). At the same time, technological changes occurring since the 1980s have resulted in the creation of capital, such as machines and computers, that can perform mainly routine tasks and can therefore substitute for workers in occupations with high routine task content. This hypothesis has become known as 'routinization', or routine-biased technical change (RBTC).

In this paper, I investigate the implications of routine-biased technical change within the context of a general equilibrium model with endogenous sorting of workers into occupations based on comparative advantage. The novel aspect of the paper is the focus on the individual-level predictions in terms of occupational switching patterns and wage changes. The paper's main contributions are to formalize these individual-level predictions within this type of model, and to test them using data from the Panel Study of Income Dynamics (PSID) from 1976 to 2007. To the best of my knowledge, the paper is the first to directly use individual-level panel data to study the labor market experience of routine workers in the U.S. over the past three decades, thus shedding light on what has happened to these workers over time.[7] The approach taken in this paper provides micro-level evidence on the dynamics underlying the aggregate patterns of employment and wage polarization, on the way in which particular subsets of workers have been impacted by routinization, and on the changes over time in occupational wage premia once selection has been accounted for.

The occupational sorting mechanism featured in the model used in this paper follows Gibbons, Katz, Lemieux, and Parent (2005): workers select into occupations based on their comparative advantage. Unlike Acemoglu and Autor (2011), and following Jung and Mercenier (2010), the model economy is composed of three distinct occupations (non-routine manual, routine, and non-routine cognitive) and a continuum of workers differentiated according to their skill level.[8] Capital is modeled as suggested by Autor, Levy, and Murnane (2003): it enters the production function as a substitute for labor working in routine tasks, and a complement for workers in non-routine cognitive tasks.

I derive the model's predictions for the effects of routine-biased technical change (RBTC). RBTC is modeled as an exogenous increase in the use of

[7] A common assumption has been that most of these workers have been displaced into low-skill jobs (e.g. Acemoglu and Autor (2011), p.64), but there has been little evidence put forth to support this claim.

[8] See also Costinot and Vogel (2010) for a model with a continuum of skills and a continuum of tasks.

physical capital (due, for example, to a fall in the cost of computing power).[9] The model makes the following predictions: RBTC induces workers at the bottom of the ability distribution within routine occupations to switch to non-routine manual jobs, while inducing those at the top end to switch to non-routine cognitive jobs. The model also makes predictions in terms of the changes in occupational wage premia: The wage premium in routine occupations is predicted to fall relative to that in the two non-routine occupational categories. For this reason, workers staying in routine jobs experience a fall in wages, relative to those staying in either non-routine manual or non-routine cognitive jobs. Within the context of the model, switchers must do at least as well as stayers in terms of wage growth. The model and its predictions can be generalized (in expectations) to a setting with two-dimensional skills.[10]

To test the predictions of the model for individual workers, the paper uses data from the Panel Study of Income Dynamics (PSID). The PSID tracks individuals over time, making it possible to document the likelihood of transitions between different types of jobs, and to analyze the wage profiles for workers with different labor market experiences. Occupations are grouped based on an aggregation of 3-digit occupation codes into the three categories used in the model: non-routine manual (service occupations), routine (sales, clerical, craftsmen, foremen, operatives, laborers), and non-routine cognitive (professional and managerial occupations).[11]

The empirical strategy involves the estimation of a wage equation derived from the model. An individual worker's potential wage in each occupation consists of an occupation-specific premium (common to all workers in the same occupation in a given year), as well as an occupation-specific return to the worker's skills. Skills are allowed to contain both observable and unobservable components. Workers select into the occupation where their potential wage is highest. The key identifying assumptions for the estimation of the wage equation are that unobservable skills are time-invariant, workers have full information about their skills, and any idiosyncratic temporary

[9] This is the view of capital that has been suggested in the literature as an explanation for employment and wage polarization (Acemoglu and Autor, 2011). See Nordhaus (2007) for evidence on the fall in the cost of computing power, and Bartel et al. (2007) on firm-level evidence on the effects of IT adoption on firms' skill requirements and human resource practices.

[10] This extension is presented in Appendix A.2. See also Yamaguchi (2012) for a model with two-dimensional skills.

[11] Full details of the occupations included in each of the categories are given in Appendix Table A.1. Section 2.7.3 discusses the robustness of the results to using an alternative classification method based directly on task data.

shocks to individual wages are independent of sectoral choice. Under these assumptions, estimating a wage equation with *occupation spell* fixed effects (interactions of individual fixed effects with occupation dummies) controls for the self-selection of workers into occupations based on unobserved ability, and allows for the consistent estimation of the changes in occupation wage premia. The estimated occupation spell fixed effects themselves are also informative, as they can be used to rank workers according to ability within occupation-year cells.[12]

Using this empirical strategy, I find that there is strong evidence of selection on ability for workers switching out of routine jobs: Low ability routine workers are more likely to switch to non-routine manual jobs, while high ability routine workers are more likely to switch to non-routine cognitive jobs. This is fully consistent with the predictions of the model.

In terms of wage growth, I find that workers staying in routine jobs perform significantly worse than workers staying in any other type of occupation. The wage premium for routine occupations is estimated to have fallen by 17% from 1976 to the mid-2000s, relative to the wage premium for non-routine manual occupations (14% when taking account of changing returns to education). Meanwhile, over the same time period, the wage premium for non-routine cognitive occupations is estimated to have risen by 25% (7% when taking account of changing returns to education) relative to the wage premium for non-routine manual occupations.

There are also significant differences in wage growth between routine workers who stay or switch to other occupations: Those who switch to non-routine manual jobs have significantly lower wage growth than stayers over short run horizons (around 14% lower over a two-year period), but subsequently recover from this losses and have significantly faster wage growth than stayers in the long run (5 to 12% higher over a 10-year period). Meanwhile, those who switch to non-routine cognitive ones have significantly higher wage growth than stayers over a variety of time horizons (6 to 12% higher over a two-year period; 14 to 16% higher over a 10-year period).

The findings in this paper contribute to the literature that studies the effects of technology on the labor market. Technology has long been thought of as a potential driver of changes in the economy's employment and wage structure. A large literature has thought of technological change as being skill-biased, in the sense that it has disproportionately favored high-skill workers (Juhn, Murphy, and Pierce (1993), Murphy and Welch (1993), Katz

[12] The empirical strategy may be extended to allow for changes over time in the return to education and the empirical results are robust to this extension. See 2.7.2.

and Autor (1999), Berman, Bound, and Machin (1998)).[13] In line with what Acemoglu and Autor (2011) call the 'canonical' model of the labor market, this literature mostly considers two types of workers (high and low skilled) performing distinct and imperfectly substitutable tasks. Empirically, the focus of this literature has been on the evolution of the college wage premium (with college education used as a proxy for skills), and the extent to which it is explained by technology through changes in the relative demand of college and non-college workers (Goldin and Katz (2008), Katz and Murphy (1992), Boudarbat, Riddell, and Lemieux (2010)), or by the intensity of capital use through capital-skill complementarities (Krusell, Ohanian, Ríos-Rull, and Violante, 2000).

The role of occupations in these types of studies was limited, as there was generally no distinction made between skills and tasks. At most, two broad occupational categories (production and non-production) were used as a proxy for skill groups (Berman, Bound, and Machin (1998), Berman, Bound, and Griliches (1994)). Recent theories of 'routinization' or routine-biased technical change (RBTC) have brought occupations and their task content to the forefront.[14] Empirical studies of the effects of RBTC for the United States have relied on repeated cross-sectional data, such as the Census or the Current Population Survey (CPS) (e.g. Autor, Katz, and Kearney (2008)), and have studied the effects of technological change on the wage structure of the economy, through changes in the occupational composition of employment.[15] Autor and Dorn (2009) study changes in employment shares across occupations for particular demographic groups, exploiting heterogeneity across metropolitan areas in the U.S. using city-level data. So far no studies have analyzed the impacts of RBTC on individual workers, which is the focus of this paper.

In contrast to the focus in the skill-biased technical change literature on

[13]See Lemieux (2008) for a survey on the evolving nature of wage inequality and the different theories that have been suggested since the 1990s.

[14]In addition to the routine content of occupations, their offshorability has also been argued to play an important role in the polarization of the labor market, as many middle-wage occupations also display task characteristics which make them more easily offshorable (Grossman and Rossi-Hansberg (2008), Firpo, Fortin, and Lemieux (2011)). Changes in the industrial composition, on the other hand, do not explain the changes in the employment structure: Acemoglu and Autor (2011) find that the shift against middle-skilled and favoring high- and low-skilled occupational categories occurs mainly within industries.

[15]Some attention has also been paid to changes in mean wages across occupations. For example, Krueger (1993) finds that occupations that have a larger increase in the share of workers using computers between 1984 and 1989 had a higher increase in mean log hourly earnings.

the changes over time in the skill wage premium, changes in occupational wage premia (implied by models of RBTC such as the one used in this paper) have not received much attention in the literature. Gibbons, Katz, Lemieux, and Parent (2005) estimate the levels of occupational wage premia, but not their changes over time. Many empirical studies include occupation dummies when estimating wage regressions, but few include occupation-year dummies within a framework that allows for their consistent estimation. For example, Cragg and Epelbaum (1996) estimate changes over time in the occupational premia in Mexico, but their estimation strategy does not take into account that selection into occupations may be correlated with workers' unobservable characteristics.[16] The empirical strategy followed in this paper, which does control for selection into occupations, generates unbiased and consistent estimates, and thus provides information on how occupational wage premia have changed in the U.S. since the mid-1970s.

The paper also contributes to the literature on occupational mobility and its associated wage changes. Kambourov and Manovskii (2009b) argue that an important component of human capital is occupation-specific, and is lost when a worker switches occupations.[17] Meanwhile Gathmann and Schonberg (2010) and Poletaev and Robinson (2008) provide evidence that human capital has an important task-specific component. Kambourov and Manovskii (2008) document an increase in occupational mobility in the United States between 1968 and 1997.[18] Groes, Kircher, and Manovskii (2009), using Danish administrative data, find a U-shaped pattern for occupational mobility (workers at the extremes of the wage distribution within an occupation are more likely to switch occupations than those in the middle), consistent with what I find in this paper. The findings in this paper bridge the gap between this literature on individual-level patterns and the aggregate-level polarization literature on the effects of technological change. The results presented in this paper help interpret many of the findings from the occupational mobility literature within the broader context of technical change and labor market polarization. Routine-biased technical change is

[16] They find that high paying occupations experienced high growth in their wage premia over the period 1987-1993, explaining close to half of the growing wage dispersion in Mexico, while low-skill occupations experienced rapid employment growth but sluggish wage growth. They link their findings to the elasticity of labor supply for workers of different skill levels.

[17] See also Sullivan (2010) for evidence on the varying degrees of importance of occupation- and industry-specific human capital across different 1-digit occupations. His analysis uses data from the 1979 Cohort of the National Longitudinal Survey of Youth (NLSY).

[18] See also Moscarini and Thomsson (2007).

theoretically consistent as the driving force behind increasing occupational mobility, selection on ability for occupational switchers, and changes in occupation wage premia over time.

The rest of the paper is organized as follows. Section 2.2 describes the theoretical framework. Section 2.3 derives the model's predictions for the effects of routine-biased technical change. Section 2.4 describes the data and the occupational categories. Section 2.5 describes the empirical strategy. Section 2.6 presents the empirical results testing the predicted effects of RBTC using PSID data. Section 2.7 presents robustness checks on the main results of the paper, and Section 2.8 concludes.

2.2 Model

The model features an economy where workers sort into occupations based on comparative advantage as in Gibbons, Katz, Lemieux, and Parent (2005).[19] There is perfect information. Following Jung and Mercenier (2010) there is a continuum of workers who are differentiated by their skill level, and three occupations.[20] Capital enters the production function as a substitute for routine tasks and a complement for non-routine cognitive tasks. Technical change is driven by increases in capital utilization (due, for example, to a fall in the cost of computing power), and is therefore routine-biased. Appendix A.2 extends the model to allow for two-dimensional skill endowments (cognitive and manual) and describes the conditions under which the predictions of the basic model are still valid in expectations in that richer model. Note that the exposition of the model in this section follows Jung and Mercenier (2010).

[19] Acemoglu and Autor (2011) call this type of model a Ricardian model of the labor market.

[20] This contrasts with Acemoglu and Autor (2011), who consider a continuum of tasks and three skill groups. One advantage of the Jung and Mercenier (2010) setup is that it does not require the definition of arbitrary distinctions between low-, middle- and high-skill workers. Boundaries only need to be defined between occupations (non-routine manual, routine, and non-routine cognitive). These distinctions can be made by relying on broad occupation codes, which differ sharply in terms of their task content (although the classification may still be subject to criticism, as some particular occupations may be hard to classify in an obvious manner). Another advantage of the setup used here is that it allows each individual worker's wage to depend both on their skill level *and* the task they perform (which as will be shown later, is empirically relevant). In Acemoglu and Autor (2011), all workers of a given skill receive the same wage, regardless of the task they are employed in.

2.2.1 Household Preferences

There is a single representative household composed of a continuum of workers. The household has Cobb-Douglas preferences over two consumption goods, Y_1 and Y_2. For reasons that will become clear later, Y_1 may be thought of as a service good, and Y_2 as a manufactured good. The household's utility function is given by:

$$U(Y_1, Y_2) = (1 - \beta) \ln Y_1 + \beta \ln Y_2 \tag{2.1}$$

where $0 < \beta < 1$. Maximizing utility subject to the budget constraint $I = p_1 Y_1 + p_2 Y_2$ (where I stands for total income) yields the following demand system:

$$p_1 Y_1 = (1 - \beta) I \tag{2.2}$$

$$p_2 Y_2 = \beta I \tag{2.3}$$

2.2.2 Firms

Both industries Y_1 and Y_2 are perfectly competitive. Y_1, the service good, is produced by labor performing non-routine manual tasks (which, in practice, are mostly service occupations). Y_2, the manufactured good, requires the combination of two different tasks: routine and non-routine cognitive. Routine tasks may be performed either by labor or by physical capital (machines, computers), while non-routine cognitive tasks may be performed only by labor. Specifically, let the production function for the Y_2-good be:

$$Y_2 = \min\{\kappa_{rt} rt, cog\} \tag{2.4}$$

where κ_{rt} are exogenously determined routine task services provided by machines (capital), rt are total routine task services provided by workers, and cog are total non-routine cognitive task services provided by workers. Thus, capital is a substitute for labor providing routine tasks, while it is a complement for labor providing non-routine cognitive tasks.[21] rt and cog are endogenous and their determination will be described in detail below.

The assumptions of perfect substitutability between capital and routine workers, and perfect complementarity between capital and non-routine cognitive workers, although admittedly extreme, capture the role of capital in

[21] See Autor et al. (2003) and Acemoglu and Autor (2011) for a discussion of why computers may be thought of as substitutes for routine workers and complements for non-routine workers.

a simple and tractable way, and allow the derivation of strong and clear predictions from the model on the effects of routine-biased technical change (as will be discussed in Section 2.3).[22]

The marginal cost of labor (the wage per efficiency unit) for each task will be determined in equilibrium and is denoted C_{man}, C_{rt} and C_{cog}, for non-routine manual, routine, and non-routine cognitive tasks, respectively.

2.2.3 Labor Productivity

Workers supply labor, and are differentiated by their skill level z, which has an exogenous cumulative distribution $G(z)$ with support $[z_{min}, z_{max}]$. Each worker may perform one of three distinct tasks: non-routine manual (*man*), routine (*rt*), or non-routine cognitive (*cog*).

Let $\varphi_j(z)$ denote the productivity (in terms of supplied efficiency units) of a worker of skill z performing task $j \in man, rt, cog$. $\varphi_j(z)$ is continuous and increasing in z so that a higher skilled worker is more productive than a less skilled one when performing the same task (absolute advantage). It is also assumed that more skilled workers have a comparative advantage in performing more complex tasks (where non-routine cognitive tasks are assumed to be more complex than routine tasks, and these in turn are assumed to be more complex than non-routine manual tasks). The productivity differences are assumed to hold not only in levels but also in logs. This means that:

$$0 < \frac{d \ln \varphi_{man}(z)}{dz} < \frac{d \ln \varphi_{rt}(z)}{dz} < \frac{d \ln \varphi_{cog}(z)}{dz} \tag{2.5}$$

Assume $\varphi_j(z_{min}) = 1$ for $j \in man, rt, cog$.

2.2.4 Worker Sorting and Wages

Workers will choose which task to perform based on the potential wage they would receive in each occupation, which is given by the competitively determined wage per efficiency unit, and the number of efficiency units supplied by the worker in that task. That is:

$$w_j(z) = C_j \varphi_j(z) \tag{2.6}$$

where $w_j(z)$ is the potential wage in occupation $j \in man, rt, cog$ for an individual of skill level z.

[22] Autor et al. (2003) use a Cobb-Douglas specification to capture the complementarity between routine and non-routine tasks, with computers being perfect substitutes for workers in providing routine tasks.

In equilibrium, workers will sort between the three types of jobs according to their respective comparative advantage (given by Equation (2.5)). In particular, there will be two endogenously determined skill thresholds z_0 and z_1 (where $z_{min} < z_0 < z_1 < z_{max}$), such that the least skilled workers, that is, those with $z \in [z_{min}, z_0)$ will find it optimal to select into the non-routine manual occupation, producing good Y_1; the medium-skill workers, that is, those with $z \in [z_0, z_1)$, will find it optimal to perform the routine task within the Y_2-sector; and the most skilled workers, i.e. those with $z \in [z_1, z_{max}]$, will find it optimal to work in non-routine cognitive jobs also within the Y_2-sector.

Wages will therefore satisfy:

$$w(z) = \begin{cases} C_{man}\varphi_{man}(z) & \text{for } z_{min} \leq z < z_0 \\ C_{rt}\varphi_{rt}(z) & \text{for } z_0 \leq z < z_1 \\ C_{cog}\varphi_{cog}(z) & \text{for } z_1 \leq z \leq z_{max} \end{cases}$$

In equilibrium, the cutoffs z_0 and z_1 are determined so that the marginal workers have no incentives to relocate between tasks. That is, the marginal worker would receive the same wage performing either task. Formally, this means:

$$C_{man}\varphi_{man}(z_0) = C_{rt}\varphi_{rt}(z_0) \tag{2.7}$$

$$C_{rt}\varphi_{rt}(z_1) = C_{cog}\varphi_{cog}(z_1) \tag{2.8}$$

According to the way in which the tasks have been labeled, this equilibrium distribution implies that mean real wages will be lowest for non-routine manual workers, and highest for non-routine cognitive workers, which is consistent with the data (as will be shown in the empirical section). Note also that an individual worker's wage depends both on his skill level, and on the type of task he performs.

2.2.5 Equilibrium

The equilibrium skill thresholds z_0 and z_1 determine the employment in each of the occupation types, and the output of each of the goods Y_1 and Y_2. For the Y_1-good, the market-clearing condition is:

$$\int_{z_{min}}^{z_0} \varphi_{man}(z)dG(z) = Y_1 \tag{2.9}$$

For the Y_2-sector, from equation (2.4), total input of routine and non-routine cognitive task services must be equal in equilibrium. That is:

$$\kappa_{rt}\int_{z_0}^{z_1}\varphi_{rt}(z)dG(z)=\int_{z_1}^{z_{max}}\varphi_{cog}(z)dG(z) \tag{2.10}$$

Recall that κ_{rt} accounts for the (exogenous) contribution of capital to the provision of routine tasks.

The market-clearing condition for the Y_2-good can be written either in terms of the total input of routine task services, or the total input of non-routine cognitive task services. In terms of the latter, it is as follows:

$$\int_{z_1}^{z_{max}}\varphi_{cog}(z)dG(z)=Y_2 \tag{2.11}$$

Marginal cost pricing holds in the Y_1-sector, so that:

$$p_1=C_{man} \tag{2.12}$$

Finally, the household's income is given by:

$$I=C_{man}\int_{z_{min}}^{z_0}\varphi_{man}(z)dG(z)+C_{rt}\kappa_{rt}\int_{z_0}^{z_1}\varphi_{rt}(z)dG(z)+C_{cog}\int_{z_1}^{z_{max}}\varphi_{cog}(z)dG(z) \tag{2.13}$$

Let the Y_1 good be the numeraire, so $p_1=1$.

Equations (2.2), (2.3), (2.7), (2.8), (2.9), (2.10), (2.11), (2.12), and (2.13) along with the choice of numeraire determine the equilibrium levels of the endogenous variables C_{man}, C_{rt}, C_{cog}, z_0, z_1, Y_1, Y_2, p_1, p_2, I.

2.3 Effects of Routine-Biased Technical Change

In this section, I analyze the effects of routinization-biased technical change (RBTC) on the endogenous variables of the model. The analysis extends Jung and Mercenier (2010) by presenting a formal derivation of the general equilibrium effects of RBTC, and by focusing on the implied effects for individual workers in terms of occupational switching patterns and wage changes. Unlike Jung and Mercenier (2010), who define RBTC as an increase in κ_{rt} as well as a simultaneous increase in the slope of $\varphi_{rt}(z)$, I define RBTC as an increase in κ_{rt} only.

I do this for two reasons. First, routinization theories have thought of capital as changing marginal productivities of workers performing different tasks due to the substitutabilities and complementarities embedded in the production function (Autor et al., 2003), rather than through changes in the supply of efficiency units of particular worker types. Although this extra channel might be worth exploring in future work, by changing only κ_{rt} I can study the implications of standard routinization theories within the context of the model considered in this paper. This ensures comparability with the predictions derived from other models in the literature, in particular, Acemoglu and Autor (2011).[23] Second, by changing only one parameter, the general equilibrium effects of that specific change may be isolated and formalized. This also eases the interpretation of the results, as all of the implied effects may be attributed to the change in one particular parameter.

2.3.1 Switching Patterns Induced by RBTC

First, consider a comparative statics analysis of the effects of a change in κ_{rt} on the ability cutoffs z_0 and z_1. This will tell us what kind of occupational switching is induced by RBTC, and which workers switch to which occupations.

Define:

$$
\begin{aligned}
man(z_0) &\equiv \int_{z_{min}}^{z_0} \varphi_{man}(z) dG(z) && (2.14)\\
rt(z_0, z_1) &\equiv \int_{z_0}^{z_1} \varphi_{rt}(z) dG(z) && (2.15)\\
cog(z_1) &\equiv \int_{z_1}^{z_{max}} \varphi_{cog}(z) dG(z) && (2.16)
\end{aligned}
$$

These are the equilibrium total labor services in non-routine manual, routine, and non-routine cognitive tasks, respectively.

Using the normalization $p_1 = 1$, and combining equations (2.2), (2.9)-(2.12), and (2.13) we can get to the following two-equation system, with two unknowns z_0 and z_1:

$$
\begin{aligned}
man(z_0) &= \frac{1-\beta}{\beta} \left[\frac{\varphi_{man}(z_0)}{\varphi_{rt}(z_0)} \left(1 + \frac{\varphi_{rt}(z_1)}{\varphi_{cog}(z_1)} \right) \right] cog(z_1) && (2.17)\\
\kappa_{rt} rt(z_0, z_1) &= cog(z_1) && (2.18)
\end{aligned}
$$

[23] Jung and Mercenier (2010) are interested in distinguishing between the effects of RBTC and the effects of globalization, and for that purpose, changing the slope of $\varphi_{rt}(z)$ as a consequence of RBTC is important.

Take logs of these equations to get:

$$\ln man(z_0) = \ln\left(\frac{1-\beta}{\beta}\right) + \alpha_0(z_0) + \tilde{\alpha}_1(z_1) + \ln cog(z_1) \quad (2.19)$$

$$\ln \kappa_{rt} + \ln rt(z_0, z_1) = \ln cog(z_1) \quad (2.20)$$

where the following definitions have been used: $\alpha_0(z_0) \equiv \ln\left[\varphi_{man}(z_0)/\varphi_{rt}(z_0)\right]$, and $\tilde{\alpha}_1(z_1) \equiv \ln\left[1 + \varphi_{rt}(z_1)/\varphi_{cog}(z_1)\right]$.

Take total derivatives of equations (2.19) and (2.20) to get:

$$\begin{pmatrix} \alpha_0'(z_0) - \frac{man'(z_0)}{man(z_0)} & \tilde{\alpha}_1'(z_1) + \frac{cog'(z_1)}{cog(z_1)} \\ -\frac{rt_0(z_0,z_1)}{rt(z_0,z_1)} & \frac{cog'(z_1)}{cog(z_1)} - \frac{rt_1(z_0,z_1)}{rt(z_0,z_1)} \end{pmatrix} \begin{pmatrix} dz_0 \\ dz_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} d\ln\kappa_{rt} \quad (2.21)$$

Proposition 1 (Effect of RBTC on ability cutoffs): The general equilibrium effects of $d\ln\kappa_{rt}$ on the cutoffs z_0 and z_1 are given by:

$$\frac{dz_0}{d\ln\kappa_{rt}} = \frac{-\tilde{\alpha}_1'(z_1) - \frac{cog'(z_1)}{cog(z_1)}}{\Delta} > 0 \quad (2.22)$$

$$\frac{dz_1}{d\ln\kappa_{rt}} = \frac{\alpha_0'(z_0) - \frac{man'(z_0)}{man(z_0)}}{\Delta} < 0 \quad (2.23)$$

where Δ is the determinant of the matrix on the left-hand-side of the system of Equation (2.21).

Proof: See Appendix A.1.1.

This proposition says that an increase in κ_{rt} (RBTC) will lead to an increase in z_0 and a decrease in z_1. This implies employment polarization: the share of routine jobs in total employment will decrease, while the share of non-routine manual and the share of non-routine cognitive jobs will increase. It also implies the following in terms of switching patterns:

Corollary 1 (Switching Patterns induced by RBTC): Let the new ability cutoffs after the change in κ_{rt} be z_0' and z_1'. An increase in κ_{rt} will lead to the following switching pattern: Workers at the bottom of the ability distribution within routine jobs, that is, those with $z \in [z_0, z_0')$, will switch to non-routine manual jobs, while workers at the top of the ability distribution within routine jobs, that is, those with $z \in (z_1', z_1)$, will switch to non-routine cognitive jobs.

Intuitively, an increase in κ_{rt} means that physical capital produces a larger amount of routine task services. Because of the technology in the Y_2-sector, physical capital and labor performing routine tasks are substitutes, while physical capital and labor performing non-routine cognitive tasks are complements. The increase in the provision of routine task services by computers induces the Y_2-sector firms to transfer workers from routine to non-routine cognitive tasks. The workers with the highest ability among the routine workers are the best suited for this switch.

With the increased household income, demand for the Y_1 good will increase as well, so the Y_1 firms will want to hire more workers to perform non-routine manual tasks. The workers with the lowest ability among the routine workers are the best suited for this switch.

2.3.2 Wage Changes Induced by RBTC

The model has predictions for wage changes, both for workers who switch occupations and for workers who stay in the same occupation. First, consider the changes induced by RBTC on the wage per efficiency unit C_j in each occupation. Because of marginal cost pricing $C_{man} = p_1$, and p_1 is normalized to 1 in any equilibrium, so C_{man} does not change. All of the wage changes should be interpreted as being relative to this normalization.

Using the comparative statics results on the effects of κ_{rt} on z_0 and z_1, along with equations (2.7) and (2.8), we have the following result:

Proposition 2 (Changes in wage per efficiency unit induced by RBTC):

$$\frac{d\ln C_{man}}{d\ln \kappa_{rt}} = 0 \qquad \frac{d\ln C_{rt}}{d\ln \kappa_{rt}} < 0 \qquad \frac{d\ln (C_{cog}/C_{rt})}{d\ln \kappa_{rt}} > 0$$

$$\frac{d\ln C_{cog}}{d\ln \kappa_{rt}} \gtreqless 0 \;\; \textit{if and only if} \;\; \alpha_1'(z_1)\left[\alpha_0'(z_0) - \frac{man'(z_0)}{man(z_0)}\right] \gtreqless \alpha_0'(z_0)\left[\tilde{\alpha}_1'(z_1) - \frac{cog'(z_1)}{cog(z_1)}\right]$$

Proof: See Appendix A.1.2.

Intuitively, from Equation (2.7), the wage per efficiency unit of routine workers relative to non-routine manual workers depends on their relative productivity at the ability cutoff z_0. Routine-biased technical change induces an increase in z_0. At a higher z, the productivity gap between routine and non-routine manual workers is greater, so the relative wage per efficiency unit of routine workers is lower.

Conversely, from Equation (2.8) the wage per efficiency unit of routine relative to non-routine cognitive workers depends on their relative productivity at the ability cutoff z_1. Routine-biased technical change induces a fall in z_1, reducing the productivity gap between routine and non-routine cognitive workers, and reducing the relative wage per efficiency unit of routine workers.

The change in the wage per efficiency unit of non-routine cognitive relative to non-routine manual workers, however, is ambiguous. The fall in z_1 tends to increase relative wages in non-routine cognitive, but the increase in z_0 works in the opposite direction. The net effect will depend on how responsive the productivity ratios are to z at the cutoff levels (given by $\alpha_0'(z_0)$ and $\alpha_1'(z_1)$) and on how much the supply of tasks change with a marginal change in the cutoffs (which in turn depends on the productivity functions and the probability density of skills at the cutoffs).[24]

The wage change induced by the shock to κ_{rt} on workers of different skill levels are described in the following proposition.

Proposition 3 (Wage changes induced by RBTC for workers of different ability levels): The wage changes induced by a positive shock to $\ln \kappa_{rt}$ are as given in the following equation, where the final column indicates each worker's occupation before and after the shock.

$$\frac{d\ln w(z)}{d\ln \kappa_{rt}} = \begin{cases} \frac{d\ln C_{man}}{d\ln \kappa_{rt}} = 0 & \text{if } z_{min} \leq z < z_0 \quad man \rightarrow man \\ \frac{d\ln C_{man}}{d\ln \kappa_{rt}} + \ln C_{man} + \ln \varphi_{man}(z) - (\ln C_{rt} + \ln \varphi_{rt}(z)) < 0 & \text{if } z_0 \leq z < z_0' \quad rt \rightarrow man \\ \frac{d\ln C_{rt}}{d\ln \kappa_{rt}} < 0 & \text{if } z_0' \leq z < z_1' \quad rt \rightarrow rt \\ \frac{d\ln C_{cog}}{d\ln \kappa_{rt}} + \ln C_{cog} + \ln \varphi_{cog}(z) - (\ln C_{rt} + \ln \varphi_{rt}(z)) \geq \frac{d\ln C_{rt}}{d\ln \kappa_{rt}} & \text{if } z_1' \leq z < z_1 \quad rt \rightarrow cog \\ \frac{d\ln C_{cog}}{d\ln \kappa_{rt}} > \frac{d\ln C_{rt}}{d\ln \kappa_{rt}} & \text{if } z_1 \leq z \leq z_{max} \quad cog \rightarrow cog \end{cases}$$

Proof: See Appendix A.1.3.

[24] This is analogous to the ambiguity in Acemoglu and Autor (2011) regarding the effect of RBTC on the wage of high skill workers relative to low skill workers.

For stayers, wage changes are given by the change in wages per efficiency unit in their particular occupation. This implies a fall in wages of stayers in routine occupations, relative to stayers in either of the non-routine occupations. Meanwhile, workers switching out of routine must do at least as well as stayers (as they always could have chosen to stay in the routine occupation).

Figure 2.1 graphically summarizes all the results on the effects of RBTC on the equilibrium skill cutoffs and wages for workers of different ability levels. The black lines in the Figure represent the original equilibrium, before the RBTC shock. The cutoff skill levels are given by z_0 and z_1. Workers with ability below z_0 optimally select into the non-routine manual occupation, those with ability above z_1 select into the non-routine cognitive occupation, and those with ability between z_0 and z_1 select into the routine occupation. Mean wages are highest in the non-routine cognitive occupation and lowest in the non-routine manual one.

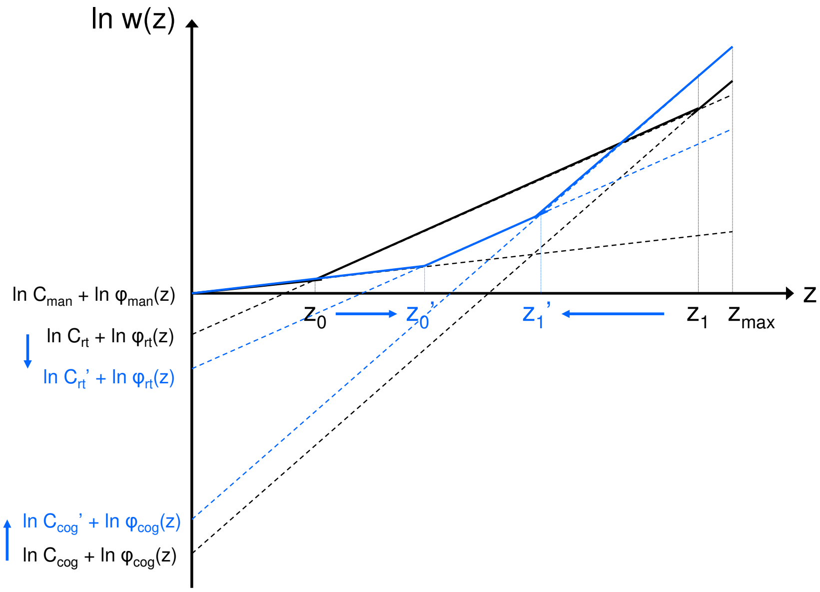
The effects of RBTC are depicted with the blue lines. From Proposition 2, the sign of the change in C_{cog} is ambiguous (although its change, if any, is greater than the change in C_{rt}). The graph is for the case where C_{cog} increases (which will prove to be empirically relevant). When the RBTC shock hits, C_{rt} falls, and workers with ability between z_0 and z_0' find it optimal to switch to non-routine manual jobs, while workers with ability between z_1' and z_1 find it optimal to switch to non-routine cognitive jobs. Stayers in routine jobs experience the largest fall in wages, given by the change in C_{rt}.

2.3.3 Summary of the Predictions of the Model

The general equilibrium effects of a positive shock to $\ln \kappa_{rt}$ are as follows:

1. Switching patterns:
    (a) The workers at the bottom of the ability distribution within routine occupations switch to non-routine manual jobs.
    (b) The workers at the top of the ability distribution within routine occupations switch to non-routine cognitive jobs.
    (c) No switching is induced for non-routine workers (either manual or cognitive).
2. Wage changes:

Figure 2.1: Effects of routine-biased technical change on skill cutoffs and wages (case where C_{cog} increases relative to C_{man})



(a) Workers staying in routine jobs experience a fall in real wages, relative to those staying in non-routine manual jobs (because C_{rt} falls).

(b) Workers staying in non-routine cognitive jobs experience an increase in real wages, relative to those staying in routine jobs (because C_{cog}/C_{rt} increases).

(c) Workers who switch from routine to non-routine jobs (either cognitive or manual) experience an increase in real wages relative to those who stay in the routine occupation.

2.4 Data

In order to test the individual-level predictions of the model, I use data from the Panel Study of Income Dynamics (PSID) for the United States. The PSID is a longitudinal study of nearly 9,000 U.S. families. Following the same families since 1968, the PSID collects data on economic, health, and social behavior, including the occupational affiliation of the household head and wife, their wage on their main job at the time of the interview, and their total labor earnings in the previous calendar year.[25] The PSID has the advantage of providing information for individuals from many different cohorts over a wide range of years. Data is available at an annual frequency between 1968 and 1997, and bi-annually from 1997 onwards.[26]

The paper uses wages reported for the current job, as they can be directly linked to the occupation that the respondent is working in at the time of the interview. Data on wages for salaried workers is only available starting in 1976, so the analysis only uses data from that year onwards.[27] The most recent data used in the paper are for 2007.

The sample is limited to male household heads, aged 16 to 64, employed in non-agricultural, non-military jobs, and who are part of the "Survey Research Center" (SRC) sample. This is the main original sample from the

[25] The Panel Study of Income Dynamics is primarily sponsored by the National Science Foundation, the National Institute of Aging, and the National Institute of Child Health and Human Development and is conducted by the University of Michigan. PSID data is publicly available at `http://psidonline.isr.umich.edu/`.

[26] Comparing trends in cross-sectional inequality across the PSID and the CPS, Heathcote, Perri, and Violante (2010) find that the two datasets track each other quite well. The only striking discrepancy is the sharp increase in the variance of CPS households earnings in the 1970s, which is not observed in the PSID.

[27] Throughout the paper, nominal values are converted to 1979 dollars using the Consumer Price Index for All Urban Consumers from the Bureau of Labor Statistics.

PSID. The over-sample of low-income households (SEO sample) and the Immigrant samples added in the 1990s are excluded from the analysis.[28]

Throughout the paper occupations are classified into three broad groups, based on the categories used by Acemoglu and Autor (2011). The groups are as follows:

- **Non-routine cognitive:** Professional, technical, management, business and financial occupations.
- **Routine:**[29] Clerical, administrative support, sales workers, craftsmen, foremen, operatives, installation, maintenance and repair occupations, production and transportation occupations, laborers.
- **Non-routine manual:** Service workers.

The categorization is based on the aggregation of 3-digit occupational codes that map into these broader categories. Each group is labeled with the name of the main task performed by workers in that occupation, as explained in Acemoglu and Autor (2011) and supported by data from the Dictionary of Occupational Titles.[30] More details on the specific occupations and occupation codes included in each category are presented in Appendix A.3. The Appendix also describes an alternative classification procedure based on task data from the Dictionary of Occupational Titles, to which the results are robust (see Section 2.7.3).

Table 2.1 presents descriptive statistics for each of the broad occupation groups. Non-routine cognitive and routine occupations account for the majority (93%) of total employment. Mean wages are highest in non-routine cognitive occupations and lowest in non-routine manual jobs. Non-routine

[28]Women are excluded as there are many confounding factors in the changes in the occupational composition of employed women over the past three decades. The low-income over-samples is excluded in order to keep a sample that is representative of the entire population (at least in the early years). The immigrant sample is excluded as it is only available for some years and significantly changes the occupational composition of employment in the sample.

[29]I do not distinguish between routine cognitive and routine manual workers in order to ensure consistency with the occupational groupings used in the model. Note that because the sample used in this paper includes only men, the vast majority of routine workers are in routine manual occupations.

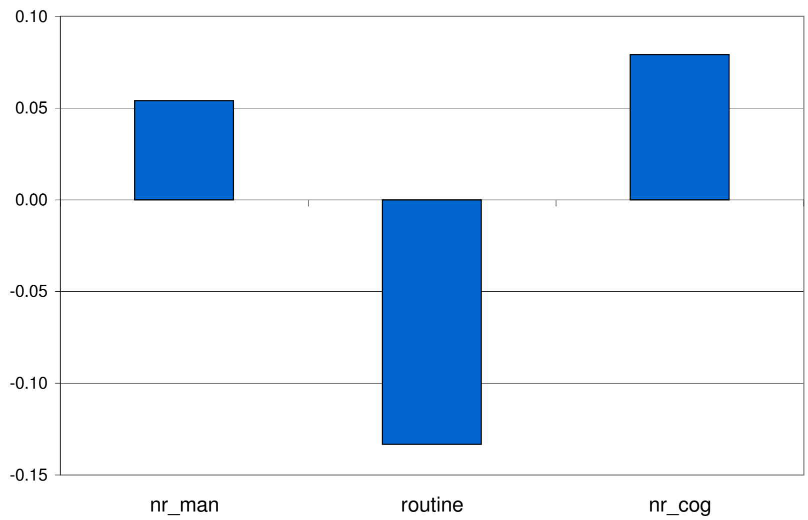
[30]In a recent working paper, Lefter and Sand (2011) suggest that the occupational coding scheme plays an important role in the extent and timing of polarization. They find that using alternative coding methods affects the extent of wage growth in low-skill occupations, and weakens the contrast between the 1980s and the 1990s in employment growth patterns, relative to what is suggested by Autor et al. (2006).

Table 2.1: Descriptive statistics (1976-2007)

	nr_cog (Prof, Manag, Tech)	routine (Prodctn, Operators, Clerical)	nr_man (Service)
Nr of Obs	21,431	25,861	3,477
Share	0.42	0.51	0.07
Averages:			
Real Wages	11.96	7.02	5.94
Age	40.67	38.24	37.40
Fractions within the occupation group:			
Union	0.07	0.25	0.26
H.S. Dropout	0.02	0.15	0.11
H.S. Grad	0.16	0.50	0.42
Some Coll.	0.22	0.24	0.32
College	0.60	0.11	0.14
Task measures (1976-2001 averages):			
DOT Non-Routine Cognitive	6.01	1.82	1.30
DOT Routine	2.97	4.50	2.40
DOT Non-Routine Manual	0.74	1.89	2.28

Note: Real wages are in 1979 dollars. Sample includes male household heads aged 16 to 64 employed in non-agricultural, non-military jobs, who are part of the PSID's core sample. The task measures are from the Dictionary of Occupational Titles (DOT) 4th Edition, published in 1977 (ICPSR, 1981). DOT task measures are aggregated to 1970 Census Occupation Codes following Autor et al. (2003). Each DOT score is rescaled to have a (potential) range from zero to 10. Following Autor et al. (2003), Non-Routine Cognitive is the mean score for 'Mathematics' and 'Direction, control and planning'; Routine is the mean score for 'Dealing with set limits, tolerances and standards' and 'Finger dexterity'; and Non-Routine Manual is the score for 'Eye-hand-foot coordination'. The averages are for the period 1976-2001, as task measures at the 1970-COC level cannot be attached to PSID data from 2003 onwards (when occupations are coded in 2000 Census codes).

Figure 2.2: Changes in employment shares for broad occupation groups, men, PSID, 1968-2007



Note: Sample includes male household heads aged 16 to 64 employed in non-agricultural, non-military jobs, who are part of the PSID's core sample. nr_cog stand for non-routine cognitive; nr_man stands for non-routine manual. See Data Section for details on the occupation classification.

cognitive jobs have a much higher share of college educated workers, and a much lower rate of unionization than the other occupations.[31] Figure 2.2 plots the long-run changes in employment shares for each of the broad occupation groups over the period 1968-2007. The pattern is broadly consistent with evidence based on Census data (Acemoglu and Autor, 2011): there is a sharp decline in the share of employment in routine occupations, with compensating share increases in both of the non-routine categories.

[31]It is worth noting, however, that the ranking of wages across the three occupation groups is not driven by the composition of workers along observable characteristics, as the same ranking is observed for the residuals from a flexible wage regression using a large number of observable individual characteristics.

2.5 Empirical Implementation

In order to test the predictions summarized in subsection 2.3.3, I estimate a wage equation and rank individuals according to their estimated ability. This section describes the empirical methodology and discusses identification issues.

From the model (Equation (2.6)), the potential wage for an individual of skill level z_i in occupation j consists of an occupation wage premium (C_j), which is common to everyone in the occupation, and on the individual's occupation-specific productivity ($\varphi_j(z_i)$). Assume that productivity is log-linear in skills; that is:

$$\ln \varphi_j(z_i) = z_i a_j \tag{2.24}$$

where a_j may be interpreted as an occupation-specific return to skills. Following Equation (2.5), assume that these occupation-specific returns are highest in the non-routine cognitive occupation and lowest in the non-routine manual one. That is:

$$a_{man} < a_{rt} < a_{cog}$$

This assumption reflects the fact that skill premia vary across occupations (see Gibbons et al. (2005)), leading to workers of different abilities self-selecting into different occupations, as described in the model.

Using the assumed functional form for productivity, and allowing for variation over time in the occupation wage premium (e.g. because of RBTC), we have the following equation for the potential wage in occupation j for individual i of skill level z_i:

$$\ln w_{ijt} = \theta_{jt} + z_i a_j \tag{2.25}$$

where i denotes the individual, j denotes the occupation, t denotes the time period, and $\theta_{jt} \equiv \ln C_{jt}$ is the occupation wage premium in occupation j at time t. Note that I am assuming that individual skills are time-invariant. This assumption will be relaxed later on to allow for certain types of time-varying skills.

The wage observed by the econometrician for individual i in period t will depend on the occupation chosen by the individual, and will be given by:

$$\ln w_{it} = \sum_j D_{ijt}\theta_{jt} + \sum_j D_{ijt} z_i a_j + \mu_{it} \tag{2.26}$$

where D_{ijt} is an occupation selection indicator which equals one if person i selects into occupation $j \in \{man, rt, cog\}$ at time t and equals zero otherwise. μ_{it} reflects classical measurement error, which is assumed to be independent of sector affiliation and therefore orthogonal to D_{ijt}. μ_{it} may also be interpreted as a temporary idiosyncratic shock that affects the wages of individual i in period t regardless of his occupational choice.

Without any restrictions to mobility, a worker will select into the occupation where he receives the highest wage. Given a fixed θ_{jt}, there will exist critical values of z_i that determine the efficient assignment of workers to occupations. Because z_i and a_j are not varying over time, and because μ_{it} is not occupation-specific, occupational mobility will be driven exclusively by changes over time in θ_{jt}.

In practice occupational mobility is not frictionless. One can think of a worker's occupational choice as being driven by z_i and θ_{jt}, as well as a noise component which is *uncorrelated* with wages. This noise component may be interpreted as a search friction, which does not affect a worker's potential wage in the different occupations, but restricts the worker from immediately selecting into his desired occupation each period. Put differently, the identifying assumption is that *conditional on the occupation fixed effects and on individual workers' skills*, selection into occupations is random (i.e. driven by a search friction that is orthogonal to skills or to any other wage determinants). Therefore, we have that in Equation (2.26): $E(\mu_{it}|\boldsymbol{D}_{ij}, z_i, \boldsymbol{\theta}_j) = 0$. An estimation procedure that controls for D_{ijt}, z_i and θ_{jt} would lead to consistent estimates.[32]

For the purposes of this paper, I am not interested in identifying a_j. Therefore, I can rewrite Equation (2.26) as:

$$\ln w_{it} = \sum_j D_{ijt}\theta_{jt} + \sum_j D_{ijt}\gamma_{ij} + \mu_{it} \tag{2.27}$$

where $\gamma_{ij} \equiv z_i a_j$. The term γ_{ij} is composed of an individual's time invariant skills and the occupation-specific returns to those skills. γ_{ij} varies for an individual across occupation spells, but stays constant whenever the individual stays in the same occupation. Equation (2.27) can be consistently estimated using fixed effects *at the occupation spell level for each individual* (that is, using a fixed effect for each individual in each occupation that they are observed in). The fixed effect effectively demeans wages for each individual within occupation spells, thus capturing the time invariant component

[32]See also Wooldridge (2002) on estimation of unbalanced panels with selection on time-invariant unobservables.

within the spell, which is precisely the unobserved effect γ_{ij}. Recall that once z_i (through the fixed effect γ_{ij}) and θ_{jt} are controlled for, selection into occupations is random (depends only on the search friction). Therefore, the regressors in Equation (2.27) are orthogonal to the mean-zero error term μ_{it} and the coefficients are consistently estimated.[33] [34]

In the empirical estimation, θ_{jt} may be captured with interactions of occupations and year dummies. The omitted category in all years will be the non-routine manual occupation (which is consistent with the model where $\theta_{man} = 0$ in any equilibrium), and all wages will be relative to this normalization. To capture changes over time that affect all occupations (including non-routine manual), and to ensure that the normalization of $\theta_{man,t}$ to zero in all years is appropriate, the estimation includes a set of aggregate year effects that are assumed to be common to all workers, regardless of their occupation or their skill level. The estimates of $\theta_{rt,t}$ and $\theta_{cog,t}$ will reflect changes in the occupation wage premium over time, due to RBTC or other shocks, relative to the base occupation. Because of the inclusion of the occupation spell fixed effects, the occupation-time fixed effects are identified only from variation over time within occupation spells. Therefore, it is necessary to normalize $\theta_{rt,t}$ and $\theta_{cog,t}$ to zero for a base year.[35] This identification argument implies that $\widehat{\theta}_{rt,t}$ and $\widehat{\theta}_{cog,t}$ should be interpreted as estimating a double difference: Rather than identifying the *level* of the occupation wage premia, they identify their *changes* over time relative to the base year, and relative to the analogous change experienced by the base occupation (non-routine manual). As the purpose of this paper is to analyze changes over time in occupational wage premia, rather than their level, these are in fact

[33] Note that although z_i includes only individual skills, in practice, the occupation spell fixed effect will capture the wage effects of *all* time-invariant characteristics of the individual that impact wages within the occupation spell, regardless of whether they reflect individual skills or other factors such as discrimination.

[34] Spell fixed effects are often used in analyses with matched employer-employee data where an outcome variables of interest is assumed to be affected by an unobserved individual component, as well as an unobserved firm component. If the objective is not to estimate the effects of any individual-specific or firm-specific variables, but rather some other time-varying variable of interest, a spell fixed effect can be used to account for all observed and unobserved time-invariant components within a job spell, and to consistently estimate the coefficient on the variable of interest. Andrews et al. (2006) describe the method. For applications, see Martins and Opromolla (2009) and Schank et al. (2007) in the context of the effect of a firm's exporter status on wages, Hummels et al. (2011) on the effects of offshoring and Gürtzgen (2009) on the relationship between firm profitability and wages. See also Harris and Sass (2010) for the use of a teacher-school 'spell' fixed effect to identify the impacts of teacher training on student achievement.

[35] I do the normalization for the initial year, 1976.

the parameters of interest.[36]

The estimation procedure also makes it possible to generate estimated occupation spell fixed effects $\widehat{\gamma}_{ij}$. They will be an estimator of the return to time-invariant skills for individual *i conditional on selecting into occupation* j. Because γ_{ij} is monotonically increasing in skill within the occupation (the coefficients on skills is common for all workers who selected into the occupation), the ranking of workers according to this measure corresponds to their ranking according to their underlying ability. In order to test the model's implications regarding switching patterns, I am only interested in a worker's relative ability *within an occupation* in a given year, so having an estimator with which I can rank workers conditional on having selected into an occupation is sufficient for my purposes.

When estimating Equation (2.27) in the data, I add an extra set of controls for marital status, unionization status, region, and a dummy for whether the individual lives in a metropolitan area (SMSA). It is assumed that these variables are orthogonal to measurement error μ_{it}, and that their return is not occupation- or skill-specific. Their inclusion will therefore not affect the consistency of the estimated coefficients.[37]

To summarize, the equation being estimated is:

$$\ln w_{it} = \sum_j D_{ijt}\theta_{jt} + \sum_j D_{ijt}\gamma_{ij} + \boldsymbol{Z'_{it}\zeta} + \mu_{it} \tag{2.28}$$

where θ_{jt} are occupation-time fixed effects, γ_{ij} are occupation spell fixed effects for each individual, and $\boldsymbol{Z_{it}}$ includes year fixed effects, marital status, unionization status, region, and SMSA. In all the estimations, standard errors are clustered at the individual level.

The empirical strategy may be extended to allow for time-varying observable skills (Section 2.7.1), as well as for changes over time in the return to observable characteristics that affect ability such as education (Section 2.7.2).

[36]Gibbons et al. (2005) analyze differences in the levels of occupational wage premia and occupational returns to skills. They estimate a quasi-differenced version of Equation (2.26) using a non-linear instrumental variables technique.

[37]It is left as future work to relax this assumption in order to allow, for example, variation in the union premium across occupations.

2.6 Results: Effects of Routine-Biased Technical Change

In this section I test the predictions of the model using the PSID data. First I present results on worker's switching patterns according to their estimated ability. Then I discuss results regarding the wage changes for workers with different occupational trajectories.

2.6.1 Switching Patterns

I begin by testing the model's implications regarding occupational switching patterns. The model predicts that RBTC induces workers at the bottom of the ability distribution within routine occupations to switch to non-routine manual jobs, and workers at the top of the distribution to switch to non-routine cognitive jobs. As discussed above, I rank routine workers according to their position in the distribution of estimated log productivity in a given year (where estimated log productivity is equal to $\widehat{\gamma}_{ij}$, the estimated occupation spell fixed effect from Equation (2.28)). Recall that γ_{ij} is monotonically increasing in underlying ability z. Therefore, I refer to the quintiles of estimated log productivity within an occupation-year as ability quintiles.

I analyze the exit rates among workers in different quintiles of the ability distribution. Figure 2.3 plots the probability of switching at each ability quintile for two different sub-samples: 1977-1989 and 1991-2005. The fraction of switchers is calculated over a two year period; that is, each bar in the graph indicates the fraction of workers from ability quintile q who switch out of routine occupations between period t and period $t+2$. Only odd years are used to generate the graph. These restrictions are imposed in order to ensure comparability with the period from 1997 onwards, when the PSID became bi-annual. The fraction of switchers is calculated over the total number of workers from each quintile who have valid occupation reports in years t and $t+2$.

The Figure shows that workers at the top of the ability distribution are more likely to switch out of routine jobs than workers of lower ability in both sub-periods. After 1991, the probability of switching increases for all ability quintiles, but particularly for the lower ability workers. This leads to a U-shaped pattern in the probability of switching after 1991, with workers at the top and the bottom of the ability distribution being more likely to switch than those in the middle.

Table 2.2 confirms that the differences across quintiles are statistically significant. The Table presents the results from a linear probability model, where the dependent variable is a dummy equal to 1 if the worker switches occupations. The regressors are a set of ability quintile dummies, with the omitted category being the middle quintile. To account for the fact that these are generated regressors, the standard errors are adjusted through a bootstrap procedure.[38] Column (1) shows that before the 1990s, workers from the top ability quintile are 8.8% more likely to switch out of routine jobs than are workers in the middle of the ability distribution. From Column (2), after 1991, both workers at the bottom and the top of the distribution are significantly more likely to switch than those in the middle.[39]

Next, I consider the direction of the switches occurring at each quintile of the ability distribution. The results are plotted in Figure 2.4. Switchers from all quintiles are more likely to go to non-routine cognitive jobs than to non-routine manual jobs. This would be expected even if the direction of switch were random, as the non-routine cognitive occupation is much larger in terms of employment than the non-routine manual one. However, there is a clear patter of selection according to ability quintiles. Consistent with the prediction of the model, the probability of switching to non-routine manual jobs is decreasing in ability, while the probability of switching to non-routine cognitive jobs is increasing in ability.[40]

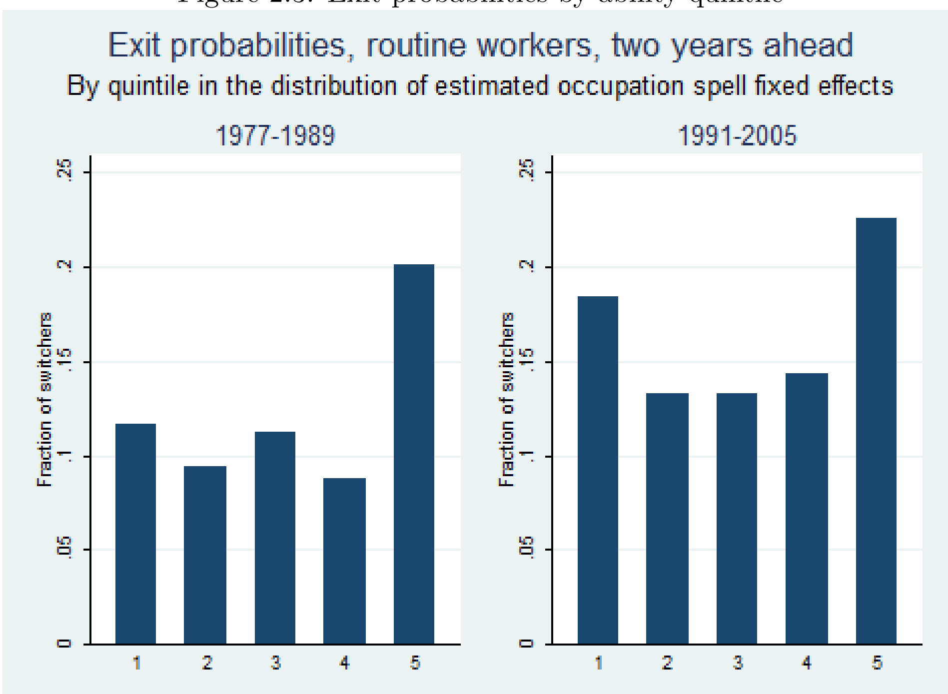
After 1990, the probability of switching to both types of non-routine occupations increases, with the probability of switching to non-routine cogni-

[38]In particular, I implement a bootstrapping procedure that performs 100 replications on randomly drawn sets of 6975 clusters of individuals (the total number available in the sample). For each randomly drawn sample, Equation (2.28) is estimated, then the estimated occupation spell fixed effects are used to rank routine workers into ability quintiles, and finally the linear probability model (with switching as the dependent variable) is estimated for these routine workers using ability quintile dummies as regressors. The standard errors presented in the Table are the bootstrapped standard errors based on these 100 replications.

[39]U-shapes in the patterns of occupational mobility have also been documented by Groes et al. (2009) using Danish administrative data. They explain these patterns within the context of a model of information frictions, where workers learn their ability level over time. This paper offers a complementary view on the reason for these U-shapes. As will be seen below, along with the U-shaped pattern of mobility observed in the PSID data, there have also been changes in the relative wage premia across occupations. Both phenomena can be explained simultaneously by the simple model of routine-biased technical change presented in this paper, while only the U-shapes are implied by the learning model in Groes et al. (2009). There is evidence, therefore, that technological change plays an important role in driving occupational mobility, although learning motives may certainly be contributing to the U-shaped patterns as well.

[40]The U-shape and the patterns in the direction of switching are also observed in the PSID data when using raw wages, or when using residuals from a flexible regression of wages on a large number of observable individual characteristics.

Figure 2.3: Exit probabilities by ability quintile



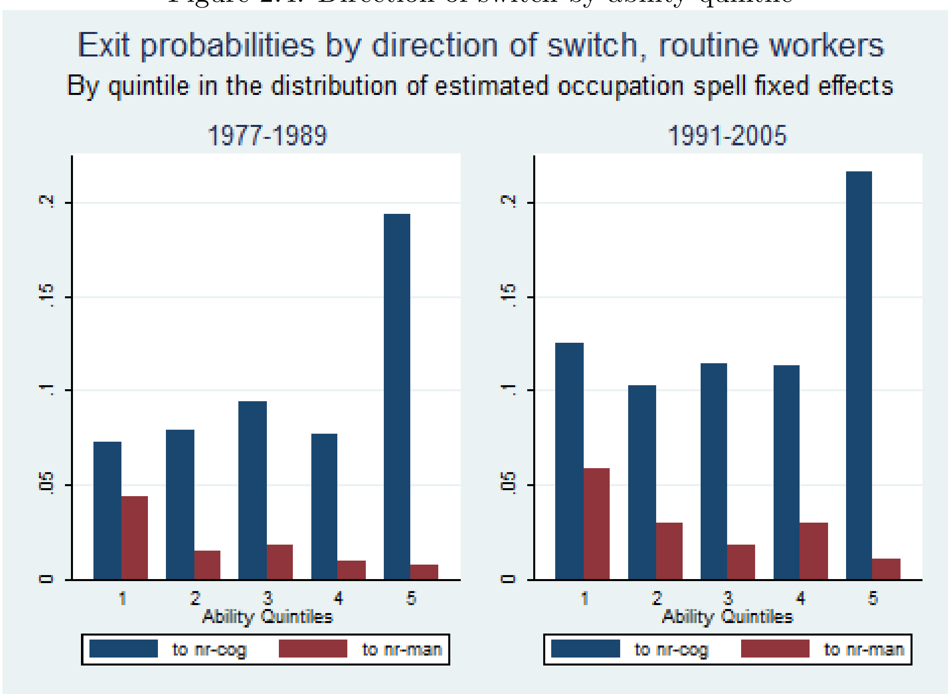
Note: Sample includes workers in routine occupations, and plots their probability of switching out of this type of occupation between years t and $t+2$, according to their ability quintile.

Table 2.2: Regressions of the probability of switching occupations two years ahead on dummies for ability quintiles (routine workers, odd years only)

	1977-1989	1991-2005
	(1)	(2)
q-1	.004	.052
	(.015)	(.015)***
q-2	-.018	.0002
	(.017)	(.014)
q-4	-.025	.011
	(.016)	(.017)
q-5	.088	.093
	(.018)***	(.016)***
Const.	.112	.132
	(.012)***	(.010)***
Obs.	5181	7512

Note: Sample includes workers in routine occupations. The dependent variable is a dummy equal to 1 if the worker is employed in a routine occupation in year t and in a non-routine occupation in year $t+2$. q-1 through q-5 represent dummies for the individual workers' estimated ability quintiles among routine workers in year t. q-1 represents the lowest ability workers and q-5 the highest ability workers. q-3 is the omitted category. Ability quintiles are built based on estimated individual-occupation fixed effects. Standard errors are bootstrapped using 100 replications based on 6975 clusters of individuals. See footnote 38 for details.

Figure 2.4: Direction of switch by ability quintile



Note: Sample includes workers in routine occupations, and plots their probability of switching to the different non-routine occupations between years t and $t+2$, according to their ability quintile.

Table 2.3: Regressions of the probability of switching to particular occupations two years ahead for routine workers (odd years only, 1977-2005)

	P(nr_cog)	P(nr_man)
	(1)	(2)
Const.	.104	.019
	(.005)***	(.002)***
post91	.030	.011
	(.006)***	(.003)***
Obs.	12693	12693

Note: Sample includes workers in routine occupations in year t. In Column (1), the dependent variable is a dummy equal to 1 if the worker is employed in a non-routine cognitive occupation in year $t+2$. In Column (2), the dependent variable is a dummy equal to 1 if the worker is employed in a non-routine manual occupation in year $t+2$. post91 is a dummy for years from 1991 onwards.

tive increasing more than the probability of switching to non-routine manual. This is documented in Table 2.3. The unconditional probability of switching to non-routine cognitive is 10.4% before 1991, and 13.4% afterwards, while the corresponding figures for non-routine manual are 1.9% and 3.0%.

Table 2.4 confirms the statistical significance of the differences across quintiles in the direction of switch patterns observed in Figure 2.4.[41] Columns (1) and (2) are linear probability models for the probability of switching to non-routine cognitive occupations for the sub-periods 1977-1989 and 1990-2005, respectively. Columns (3) and (4) are analogous regressions for the probability of switching to non-routine manual. Workers in the middle of the ability distribution (third quintile) are the omitted category. Columns (1) and (2) show positive and significant coefficients on the dummies for quintile 5, meaning that high ability routine workers are significantly more likely to go to non-routine cognitive occupations than those in the middle of the distribution. Meanwhile, workers in the bottom quintile are significantly more likely to switch to non-routine manual occupations than those in the middle, as evidenced by the findings in Columns (3) and (4).

A common concern with datasets such as the PSID is the prevalence of coding error in the occupational affiliation data (see Kambourov and Manovskii (2004) and Kambourov and Manovskii (2008)). One might be concerned, for example, that the workers at the top and the bottom of

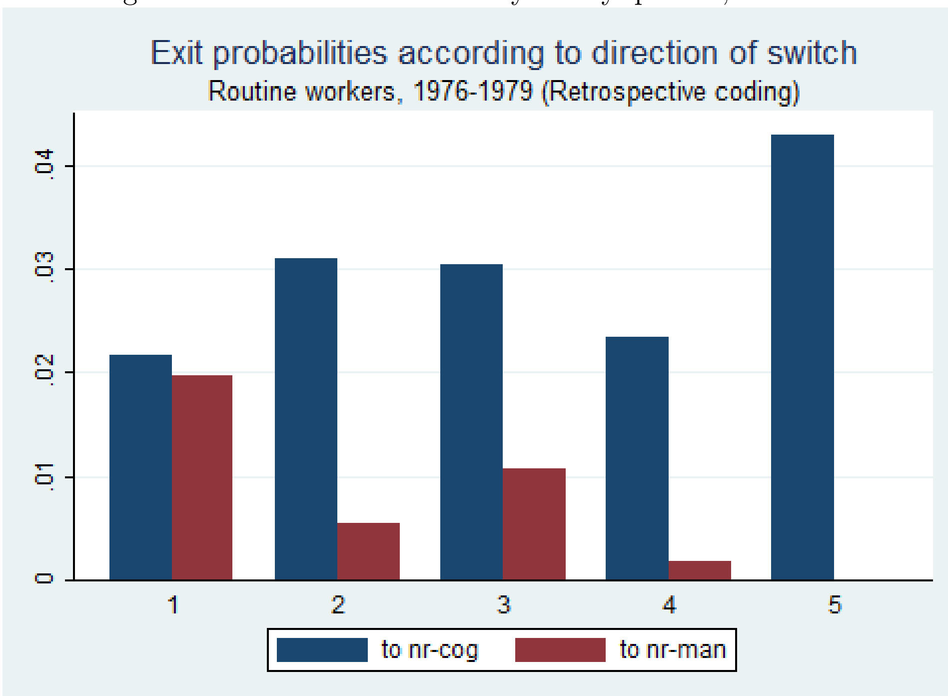
[41] The standard errors in Table 2.4 are obtained through the same bootstrap procedure as those in Table 2.2. See footnote 38.

Table 2.4: Regressions of the probability of switching to particular occupations two years ahead for routine workers (odd years only)

	P(nr_cog)	P(nr_cog)	P(nr_man)	P(nr_man)
Sub-period:	1977-1989	1990-2005	1977-1989	1990-2005
	(1)	(2)	(3)	(4)
q-1	-.022	.011	.026	.040
	(.014)	(.012)	(.007)***	(.009)***
q-2	-.015	-.011	-.002	.011
	(.017)	(.013)	(.006)	(.006)*
q-4	-.017	-.0008	-.008	.012
	(.015)	(.015)	(.005)	(.006)**
q-5	.099	.101	-.010	-.008
	(.018)***	(.015)***	(.005)**	(.005)*
Const.	.095	.114	.018	.018
	(.011)***	(.009)***	(.005)***	(.004)***
Obs.	5181	7512	5181	7512

Note: Sample includes workers in routine occupations in year t. In Columns (1) and (2), the dependent variable is a dummy equal to 1 if the worker is employed in a non-routine cognitive occupation in year $t+2$. In Columns (3) and (4), the dependent variable is a dummy equal to 1 if the worker is employed in a non-routine manual occupation in year $t+2$. q-1 through q-5 represent dummies for the individual workers' estimated ability quintiles among routine workers in year t (with q-1 representing the lowest ability workers and q-5 the highest). Workers in the middle of the ability distribution (q-3) are the omitted category. Ability quintiles are based on estimated individual-occupation fixed effects. Standard errors are bootstrapped using 100 replications based on 6975 clusters of individuals. See footnote 38 for details.

Figure 2.5: Direction of switch by ability quintile, 1976-1979



Note: Sample includes workers in routine occupations for the years 1976-1979, and plots their probability of switching to the different non-routine occupations between years t and $t+1$, according to their ability quintile.

the ability distribution within routine jobs might actually be non-routine workers who are miscoded, and thus that their transitions are spurious. To address this concern I analyze the data for the period up to 1980, when the occupational coding was done retrospectively and the prevalence of errors is thus far less severe (see Kambourov and Manovskii (2008) for full details on this argument). The results, using transitions over 1-year horizons, are shown in Figure 2.5. The patterns in the direction of switch are present even for this period, with high ability routine workers being more likely to switch to non-routine cognitive jobs, and low ability routine workers being relatively more likely to switch to non-routine manual jobs. The Figure confirms that the results are not driven by coding error. Overall, the findings on switching patterns support the predictions of the model.

2.6.2 Wage Changes

The next step is to explore the behavior of wages and wage changes. I begin with a simple motivating analysis to determine whether an individual's occupation at time t has explanatory power over his subsequent wage growth. Table 2.5 shows the results of a regression of individual wage growth between periods t and $t+j$ (where j ranges from 1 to 20 years) on dummies for the individual's occupation at time t. All regressions include year dummies. In all cases, workers in non-routine manual occupations in year t are the omitted category.

The Table shows that individuals who start a given period in a routine job have significantly lower wage growth over subsequent years than workers in non-routine occupations. This is true over time horizons as long as 20 years. For example, a worker holding a routine job in a given year can expect his real wages to grow on average 6.2% less over the subsequent four years than workers in other occupations, regardless of his future job transitions. The next sub-sections separately analyze the wage changes for stayers in routine jobs and for switchers, and take into account heterogeneities across individuals. This allows a comparison of the data with the predictions of the model.

Wage changes for occupation stayers

Consider first the wage changes for workers who do not switch occupations. Table 2.6 shows the results from running the same regressions as in Table 2.5, but now the sample includes only occupation stayers. In Column (1), stayers are defined as workers who are observed in the same broad occupation in years t and $t+1$, while in the remaining columns they are defined as workers who are observed in the same broad occupation in years t and $t+2$ (even though the wage changes may be taken over horizons longer than two years). The last two columns split the sample into two different sub-periods, 1976-1989 and 1990-2005, and consider wage changes over two-year horizons within those sub-periods. In all cases, those workers who are classified as stayers in non-routine manual jobs are the omitted category.

The Table shows that those who stay in routine jobs have significantly lower wage growth than stayers in any of the other occupational categories. For example, a worker staying in a routine job over the course of two years has a wage growth that is 1.8% lower than that of a worker remaining in a non-routine manual job over the same time period. Note that the rate of wage growth for routine workers is also in all cases significantly lower than

Table 2.5: Regression of changes in log real wages over different time horizons on dummies for initial occupation

	Change in log real wages between year t and year:					
	$t+1$	$t+2$	$t+4$	$t+10$	$t+15$	$t+20$
	(1)	(2)	(3)	(4)	(5)	(6)
nr-cog	-.016	-.017	-.013	.006	.013	.045
	(.006)***	(.006)***	(.010)	(.023)	(.034)	(.060)
routine	-.029	-.041	-.062	-.105	-.166	-.170
	(.006)***	(.006)***	(.009)***	(.022)***	(.032)***	(.059)***
Const.	.044	.121	.078	.103	.153	.216
	(.008)***	(.010)***	(.013)***	(.025)***	(.035)***	(.060)***
Obs.	31328	37114	30255	16072	8433	3752
# of Indiv.	3855	4764	4129	2756	1848	1225
R^2	.014	.02	.026	.046	.065	.062

Note: nr-cog is a dummy equal to 1 if the individual is employed in a non-routine cognitive occupation at time t. routine is a dummy equal to 1 if the individual is employed in a routine occupation at time t. Workers employed in a non-routine manual occupation at time t are the omitted category. All regressions include year dummies. Observations with log real hourly wages below 0.1 ($1.1 1979 dollars) or above 4 ($54.6 1979 dollars) excluded. Standard errors are clustered at the individual level. Due to data constraints, Column (1) uses data only up to 1997.

Table 2.6: Regression of changes in log real wages over different time horizons on dummies for initial occupation for workers who do not switch occupations

	Change in log real wages between year t and year:					
	$t+1$	$t+2$	$t+4$	$t+10$	$t+2$	$t+2$
Period:	1976-1997	1976-2007	1976-2007	1976-2007	1976-1989	1990-2005
	(1)	(2)	(3)	(4)	(5)	(6)
nr-cog	.010	.019	.027	.060	.010	.027
	(.004)**	(.005)***	(.009)***	(.023)***	(.008)	(.006)***
routine	-.009	-.018	-.038	-.079	-.024	-.014
	(.004)**	(.005)***	(.008)***	(.022)***	(.007)***	(.006)**
Const.	.020	.094	.043	.069	.101	.115
	(.007)***	(.009)***	(.012)***	(.025)***	(.011)***	(.009)***
Obs.	28029	31930	25389	13439	15999	15931
# of Indiv.	3750	4518	3832	2540	2650	3538
R^2	.016	.026	.034	.057	.025	.022

Note: Workers staying in non-routine manual occupations are the omitted category. All regressions include year dummies. The wage changes are taken over the time horizons indicated above each column (in years). The sample includes only occupational stayers. For column (1), stayers are defined as workers in the same broad occupation in years t and $t+1$. For column (2) onwards, stayers are defined as workers in the same broad occupation in years t and $t+2$ (even though the wage change may be taken over a longer horizon). Observations with log real hourly wages below 0.1 ($1.1 1979 dollars) or above 4 ($54.6 1979 dollars) excluded. Standard errors are clustered at the individual level.

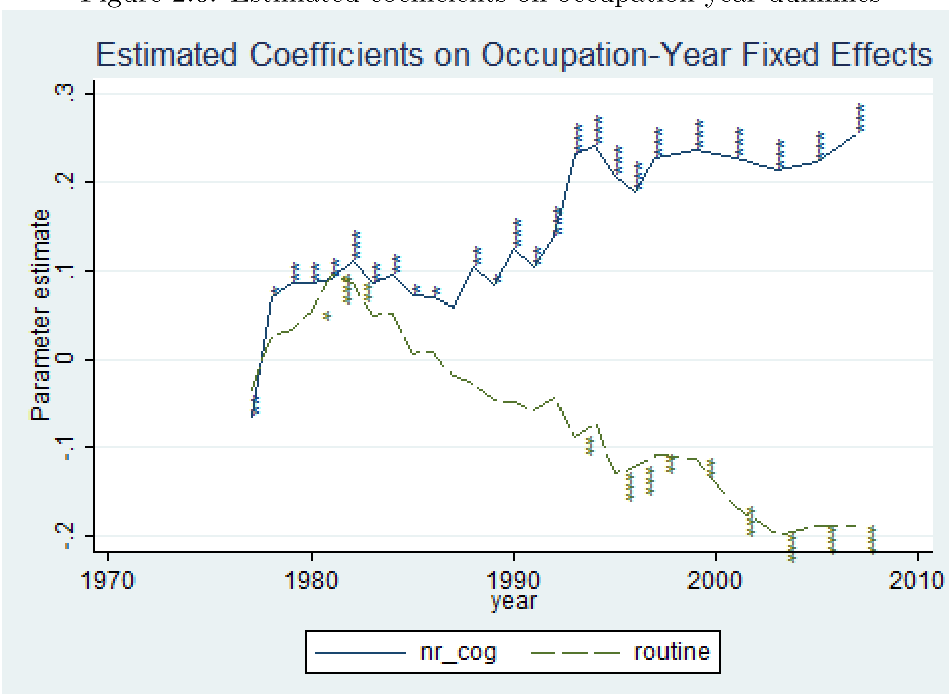
that of non-routine *cognitive* workers.

From Proposition 3, the wage changes for stayers are given by the changes in C_j for their respective occupation. Empirically, from the estimation of Equation (2.27), the estimated $\widehat{\theta}_{jt}$ (occupation-time fixed effects) will track changes in C_j over time. The Proposition implies that θ_{rt} should fall over time (relative to the omitted category). The Proposition is ambiguous about the trend of θ_{cog} relative to the omitted category, but does predict that θ_{cog} increases relative to θ_{rt}. Figure 2.6 plots the estimates of $\widehat{\theta}_{rt}$ and $\widehat{\theta}_{cog}$.

The figure shows that from the early 1980s onwards, the estimated fixed effects for routine occupations have a clear downward trend. Meanwhile, the corresponding fixed effects for non-routine cognitive occupations show an upward trend, particularly from the 1990s onwards.[42] Note that all of

[42] This is consistent with Acemoglu and Autor (2011), who talk about two different 'eras'

Figure 2.6: Estimated coefficients on occupation-year dummies



Note: The figure shows the estimates of $\widehat{\theta}_{rt}$ and $\widehat{\theta}_{cog}$ obtained from the wage equation (2.28). Stars denote the level at which the estimated coefficients are significantly different from zero.

the coefficients for the latter periods are significantly different from zero. This means that the data agree with the predictions of the model for wage stayers: Wages fall significantly for stayers in routine occupations, relative to stayers in either of the non-routine categories. The data also show a significant increase in the wage for stayers in non-routine cognitive relative to stayers in non-routine manual. Note also that the magnitude of the fall in the occupation wage premium for routine jobs is substantial. The fall from its peak in the early 1980s until the mid 2000s is similar in magnitude to the estimated rise in the college wage premium over that period.[43]

Wage growth according to direction of switch

Next I study the wage changes for routine workers who follow different switching patterns. Table 2.7 restricts the sample to routine workers only (both stayers and switchers). The dependent variable is the wage change, and the regressors are dummies for the direction of occupational switching (either to non-routine cognitive or to non-routine manual). Staying in routine jobs is the omitted category. The estimated coefficients reflect the differential wage growth for each type of switcher, relative to the stayers. Column (1) defines switchers and stayers based on individuals' occupational codes in years t and $t+1$, while the remaining columns are based on the codes in years t and $t+2$.

Panel A uses changes in real wages, while Panel B uses changes in fitted model wages, that is, changes over time in $\widehat{\theta}_{jt} + \widehat{\gamma}_{ij}$. For reference purposes, Panel C reports the percentage of routine workers classified into each of the switching categories.

The Table shows significantly lower wage growth for switchers to non-routine manual over horizons up to two years, both for real and for fitted model wages. This negative differential, however, goes away when considering longer horizons (10 years), becoming positive and significant. For example, when using fitted model wages, workers switching from a routine job in year t to a non-routine manual job in year $t+2$ experience a wage change that is 14% lower than that experienced by stayers in routine jobs.

in the changes in the distribution of wages: 1974-1988 and 1988-2008. During the first period, earnings increased monotonically with the percentile in the earnings distribution. During the second period, in contrast, growth of wages by percentiles is polarized, or U-shaped. The U-shape is more pronounced during the period 1988-1999. For employment shares, they also document a U-shaped pattern during the 1990s.

[43]Changes in the college wage premium will be discussed in further detail below in Section 2.7.2 and Figure 2.8.

By year $t+10$ however, the wage change for these workers is 5% above that of stayers.

Over all time horizons, those who switch to non-routine cognitive have significantly faster wage growth than stayers. Fitted model wages grow 12% faster over a two-year period for switchers to non-routine cognitive occupations, relative to those who stay in routine jobs. The figure is similar (14%) over a 10 year horizon.

Columns (5) and (6) in the Table show interesting differences between the periods before and after 1990. The wage gains for those who switch to non-routine cognitive are substantially larger after 1990 (18% above stayers in terms of fitted model wages after 1990, relative to 5% in the earlier period), while the wage cuts for those who switch to non-routine manual are somewhat smaller in magnitude (13% below stayers in terms of fitted model wages after 1990, relative to 15% in the earlier period).

One potential concern with the results in Table 2.7 is that the sample of workers included in each column varies according to the availability of the data, and this may be biasing the results. To address this concern, I run the regressions for changes in log real wages from Table 2.7 keeping the same set of workers over the different time horizons (that is, I only keep workers for which I have data at t, $t+2$, $t+4$ *and* $t+10$). The results are presented in Columns (1) through (3) in Table 2.8 and are very similar to the results in the previous Table. Therefore, the results in Table 2.7 are not driven by differential attrition across switchers and stayers.

So far the regressions presented have considered the implications of occupational switches in the *short-run* (between years t and $t+2$) on individual workers' wage changes *both in the short- and long-run* (between t and $t+2$, $t+4$ and $t+10$). An interesting question is whether the workers who switch out of routine occupations in the short run remain in their new occupation in future periods, or whether their subsequent switching patterns explain the long-run wage changes observed in the data. For example, if workers who switch to non-routine manual jobs in the short-run subsequently switch to other occupations in the long-run, this might be driving the finding that their wage growth is slower in the short run but faster in the long run. A full study of the occupational histories for different workers is beyond the scope of this paper (and would be difficult to perform using PSID data, given its small sample size). However, to explore whether subsequent switching patterns might be a concern, I repeat the analysis from Table 2.7, including only workers who are still observed in their $t+2$ occupation in subsequent years. That is, workers classified as stayers (switchers) will be those who stay in the routine occupation (switch to non-routine) between years t and

Table 2.7: Wage changes for routine workers, according to direction of switch

Panel A: Dependent variable is change in log real wages

	Change in log real wages between year t and year:					
	$t+1$	$t+2$	$t+4$	$t+10$	$t+2$	$t+2$
Period:	1976-1997	1976-2007	1976-2007	1976-2007	1976-1989	1990-2007
	(1)	(2)	(3)	(4)	(5)	(6)
goto-nrcog	.034	.059	.085	.163	.024	.087
	(.008)***	(.008)***	(.010)***	(.019)***	(.013)*	(.011)***
goto-nrman	-.112	-.143	-.035	.115	-.162	-.128
	(.023)***	(.023)***	(.026)	(.046)**	(.033)***	(.031)***
Const.	.037	.066	.016	-.002	.067	.041
	(.007)***	(.009)***	(.011)	(.018)	(.009)***	(.010)***
Obs.	15800	18341	14278	7568	9364	8977
# of Indiv.	2655	3253	2701	1735	1810	2425
R^2	.013	.028	.033	.061	.025	.025

Panel B: Dependent variable is change in fitted model wages (in logs)

	Change in fitted model wages between year t and year:					
	$t+1$	$t+2$	$t+4$	$t+10$	$t+2$	$t+2$
Period:	1976-1997	1976-2007	1976-2007	1976-2007	1976-1989	1990-2005
	(1)	(2)	(3)	(4)	(5)	(6)
goto-nrcog	.086	.122	.098	.139	.050	.180
	(.010)***	(.009)***	(.008)***	(.011)***	(.012)***	(.011)***
goto-nrman	-.152	-.139	-.030	.054	-.151	-.128
	(.023)***	(.021)***	(.019)	(.027)**	(.032)***	(.027)***
Const.	-.038	.026	.049	-.014	.029	-.034
	(.002)***	(.003)***	(.004)***	(.008)*	(.002)***	(.004)***
Obs.	15800	18341	14278	7568	9364	8977
# of Indiv.	2655	3253	2701	1735	1810	2425
R^2	.168	.174	.147	.09	.158	.22

Panel C: Fraction of routine workers in each of the switching categories (%)

	Fraction of routine workers in year t switching to non-routine jobs in year:					
	$t+1$	$t+2$	$t+2$	$t+2$	$t+2$	$t+2$
Period:	1976-1997	1976-2007	1976-2007	1976-2007	1976-1989	1990-2005
	(1)	(2)	(3)	(4)	(5)	(6)
goto-nrcog	8.07	10.95	11.26	11.47	9.43	12.54
goto-nrman	1.51	2.18	1.92	1.88	1.83	2.55

Note: Workers who stay in routine occupations are the omitted category. All regressions include year dummies. The wage changes are taken over the time horizons indicated above each column (in years). For column 1, occupation transitions between years t and $t+1$ are considered. For column 2 onwards, occupation transitions between years t and $t+2$ are considered (even though the wage change may be taken over a longer horizon). Observations with log real hourly wages below 0.1 ($1.1 1979 dollars) or above 4 ($54.6 1979 dollars) excluded. Standard errors are clustered at the individual level.

$t + 2$ *and* are still in the routine (non-routine) occupation in the longer run (i.e. in year $t+4$ or $t+10$). People who switch occupations between $t+2$ and the later years are dropped from the sample.[44] The results are presented in Columns (4) and (5) of Table 2.8 and confirm the main findings: Switchers to non-routine cognitive experience faster wage growth than stayers over a variety of time horizons. The same is true for the long-run wage performance of switchers to non-routine manual occupations. In short, the findings on wage changes for switchers provide support for the predictions of the model regarding the effects of RBTC.

2.7 Robustness Checks

This section presents a set of robustness checks on the empirical results of the paper.

2.7.1 Time-Varying Skills

The empirical strategy may be extended to allow for time-varying skills, under the maintained assumption that *unobservable* skills are time-invariant. To do this, Equation (2.24) can be modified to allow z_i to be a vector composed of two types of variables: a set of time-varying characteristics X_{it}, and a set of fixed characteristics η_i, each of which has an occupation-specific return which is fixed over time. That is:

$$\ln \varphi_j(\boldsymbol{z_{it}}) = X_{it}\beta_j + \eta_i b_j \tag{2.29}$$

Following Equation (2.5), assume that $\ln \varphi_j(\boldsymbol{z_{it}})$ is increasing in both of

[44] Note that this does not condition on the full occupational history, only on the $t + 2$ and either $t + 4$ or $t + 10$ occupation. Note also that an important fraction of workers do in fact switch occupations between $t + 2$ and the longer horizons, so the sample size for these regressions is smaller than in Table 2.7. For example, based on the sample in Table 2.7, of those switching from routine to non-routine cognitive between t and $t + 2$, 38% switch again between $t + 2$ and $t + 4$ and 43% do so between $t + 2$ and $t + 10$ (mostly back to routine). Of those switching from routine to non-routine manual between t and $t + 2$, 49% switch again between $t+2$ and $t+4$ and 62% do so between $t+2$ and $t+10$. Of those staying in routine between t and $t+2$, 10% switch between $t+2$ and $t+4$ and 19% switch between $t + 2$ and $t + 10$. This evidences a large degree of measured churning, even at the level of these very broad occupational categories. See also Kambourov and Manovskii (2008) for evidence of high mobility rates across 1-digit occupational categories.

Table 2.8: Robustness checks for the wage changes for routine workers, according to direction of switch

	Change in log real wages between year t and year:				
	$t+2$	$t+4$	$t+10$	$t+4$	$t+10$
	(1)	(2)	(3)	(4)	(5)
to nr-cog	.045	.092	.170	.129	.282
	(.012)***	(.014)***	(.020)***	(.014)***	(.026)***
to nr-man	-.124	.011	.112	-.083	.160
	(.038)***	(.035)	(.047)**	(.042)**	(.065)**
Const.	.095	.032	.016	.014	.013
	(.013)***	(.013)**	(.019)	(.011)	(.019)
Obs.	6553	6553	6553	12435	5953
e(N-clust)	1496	1496	1496	2515	1520
R^2	.033	.038	.061	.04	.095

Note: Workers who stay in routine occupations are the omitted category. All regressions include year dummies. The wage changes are taken over the time horizons indicated above each column (in years). For columns (1) through (3), occupation transitions between years t and $t+2$ are considered (even though the wage change may be taken over a longer horizon). In columns (4) and (5), only workers who are still in their $t+2$ occupation in the terminal year ($t+4$ in column (4), $t+10$ in column (5)) are included in the sample. Observations with log real hourly wages below 0.1 ($1.1 1979 dollars) or above 4 ($54.6 1979 dollars) excluded. Standard errors are clustered at the individual level.

its arguments. That is:

$$\beta_{man} < \beta_{rt} < \beta_{cog}$$
$$b_{man} < b_{rt} < b_{cog}$$

Observed wages for individual i in period t are then given by:

$$\ln w_{it} = \sum_j D_{ijt}\theta_{jt} + \sum_j D_{ijt}X_{it}\beta_j + \sum_j D_{ijt}\eta_i b_j + \mu_{it} \tag{2.30}$$

where, assuming (as before) that μ_{it} is independent of sector affiliation, one can now think of a worker's occupational choice as being driven by η_i, X_{it}, and θ_{jt}, as well as a noise component which is uncorrelated with wages (search friction). Thus, we have that: $E(\mu_{it}|\boldsymbol{D}_{ij}, \boldsymbol{X}_i, \eta_i, \boldsymbol{\theta}_j) = 0$. Following the same logic used to obtain Equation (2.28), I can rewrite Equation (2.30) as:

$$\ln w_{it} = \sum_j D_{ijt}\theta_{jt} + \sum_j D_{ijt}X_{it}\beta_j + \sum_j D_{ijt}\gamma_{ij} + \boldsymbol{Z'_{it}\zeta} + \mu_{it} \tag{2.31}$$

Assuming that all unobservable skills are time-invariant, Equation (2.31) can be empirically estimated using occupation-year and occupation spell fixed effects, as well as controls for (observable) time-varying skills and their occupation-specific returns. In practice, the strongest candidate variable to include in X_{it} is total work experience, which may be proxied by age.[45] I estimate Equation (2.31) including a set of dummies for 10-year age bins interacted with occupation dummies.[46] The resulting estimated occupation-year fixed effects are presented in Figure 2.7. The fall in the occupation wage premium in routine jobs remains significant and is very close in magnitude to that in Figure 2.6.

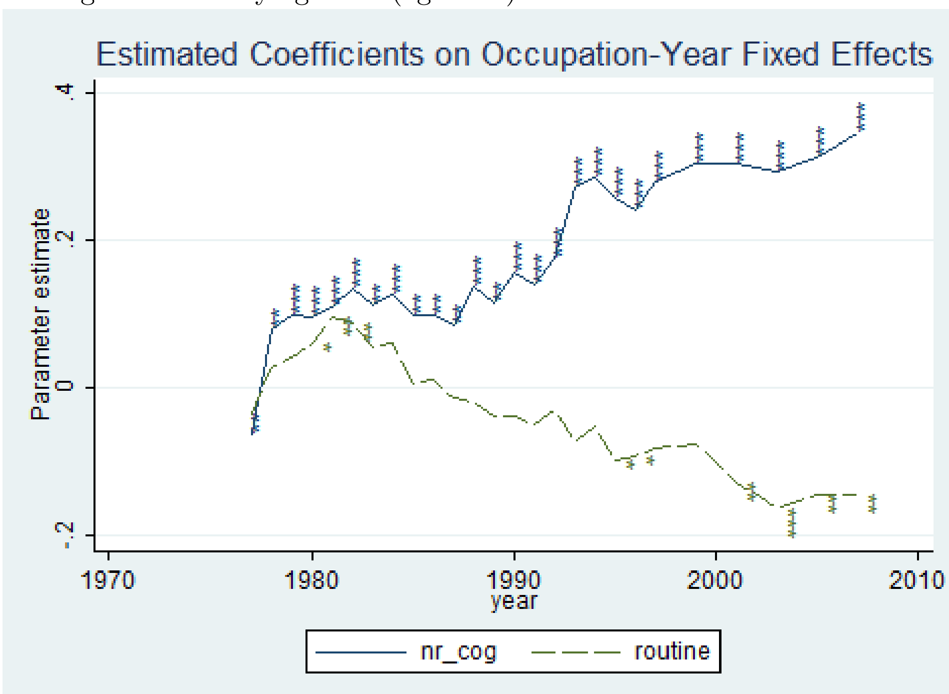
2.7.2 Changing Returns to Education

The empirical strategy may also be extended to allow changes over time in the *return* to observable characteristics that affect ability. Specifically, there is evidence that the college premium has changed over the past four decades

[45] Note that X_{it} may only include time-varying variables that reflect *general* and not occupation-specific ability, as the framework assumes that X_{it} is fully transferable between occupations (although its return varies across occupations).

[46] The bins are for age below 25, 25-34, 35-44, 45-54, and 55 and over. Because the estimation includes both occupation spell and time fixed effects, and age increases linearly over time, it is not possible to control for age directly, but only through age bins.

Figure 2.7: Estimated coefficients on occupation-year dummies when controlling for time-varying skills (age bins)



Note: The figure shows the estimates of $\widehat{\theta}_{rt}$ and $\widehat{\theta}_{cog}$ obtained from the wage equation (2.31). Stars denote the level at which the estimated coefficients are significantly different from zero.

in the United States (see for example Goldin and Katz (2008) and Acemoglu and Autor (2011)). To account for this, assume that, in Equation (2.29) all individual skills are fixed, but the return to certain kinds of (observable) skills is allowed to vary over time. That is, rewrite Equation (2.29) as:

$$\ln \varphi_{jt}(\boldsymbol{z_i}) = X_i \beta_{jt} + \eta_i b_j \tag{2.32}$$

Following Equation (2.5), assume that, for all t:

$$\beta_{man,t} < \beta_{rt,t} < \beta_{cog,t}$$
$$b_{man} < b_{rt} < b_{cog}$$

Think of X_i as education and η_i as all other individual skills.[47] The maintained assumption is that b_j, the return to all other skills, is not time-varying. For simplicity, assume that the time variation in the return to education is the same for all occupations; that is: $\beta_{jt} = \beta_j + \beta_t$. Then, the potential wages for individual i in occupation j at time t would be given by:

$$\ln w_{ijt} = \theta_{jt} + X_i \beta_j + X_i \beta_t + \eta_i b_j + \boldsymbol{Z'_{it}\zeta} + \epsilon_{it} \tag{2.33}$$

where ϵ_{it} is measurement error, which as before is orthogonal to education, ability and the wage premia, and is independent of sector affiliation. The following equation can be estimated using occupation spell fixed effects:

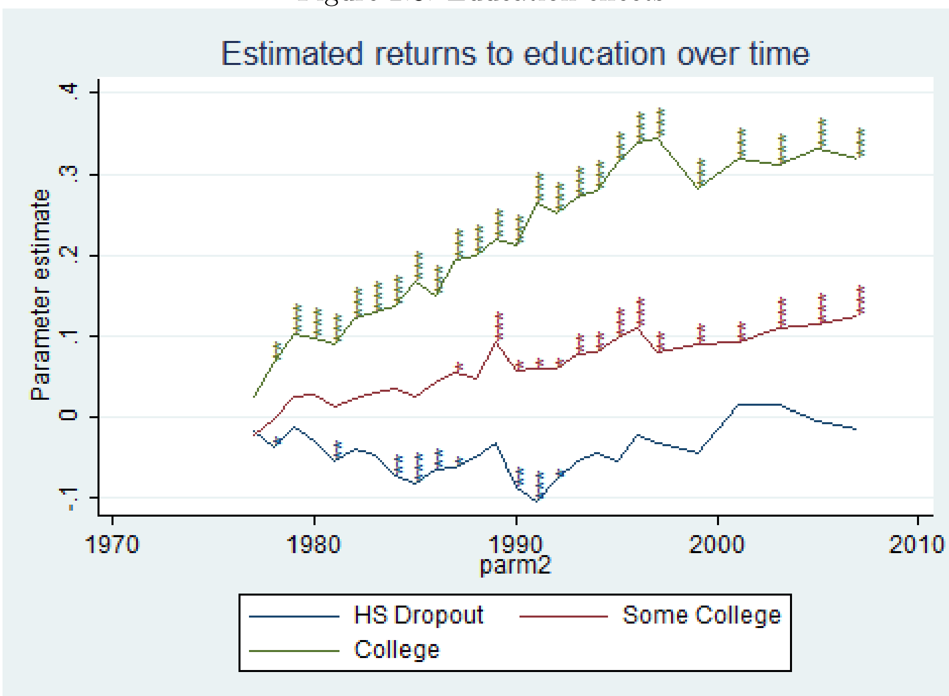
$$\ln w_{it} = \sum_j D_{ijt} \theta_{jt} + X_i \beta_t + \sum_j D_{ijt} \nu_{ij} + \boldsymbol{Z'_{it}\zeta} + \epsilon_{it} \tag{2.34}$$

where ν_{ij} are the occupation spell fixed effects, and they are such that $\nu_{ij} \equiv X_i \beta_j + \eta_i b_j$. The estimated occupation spell fixed effects and the estimated occupation wage premia will now be purged of the time-varying return to education. The occupation spell fixed effect will now only include the return to education in the base year, and the return to unobserved ability. Rankings on ability can now be constructed based on $X_i \widehat{\beta}_t + \widehat{\nu}_{ij}$.

I classify individuals into four occupation groups: high school dropouts, high school graduates, some college, and college graduates. The estimation strategy allows identification of changes in the return to education relative to a base year, and relative to the analogous change experienced by an excluded education category. Figure 2.8 shows the estimated returns to education for high school dropouts, workers with some college education,

[47] Note that the PSID does not ask individuals their education level every year. Therefore, I assign each individual their highest reported education level in any survey year, making each individual's level of education fixed over time.

Figure 2.8: Education effects



Note: Estimated coefficients are obtained from the estimation of the wage equation (2.34). Stars denote the level at which the estimated coefficients are significantly different from zero.

and college graduates, relative to the base year (1976) and relative to the omitted education category (high school graduates). The Figure confirms the finding in the literature that there has been an important rise in the return to college degrees, particularly during the 1980s and up to the mid-1990s.

Table 2.9 shows the results for switching patterns according to ability quintiles, and confirms the findings from the main body of the paper: workers in the middle of the distribution are less likely to switch than those at the extremes, and there is selection on ability in the direction of switching.[48]

Figure 2.9 plots the new estimated occupation-year fixed effects. Changing returns to education account for a sizable portion of the increase in the

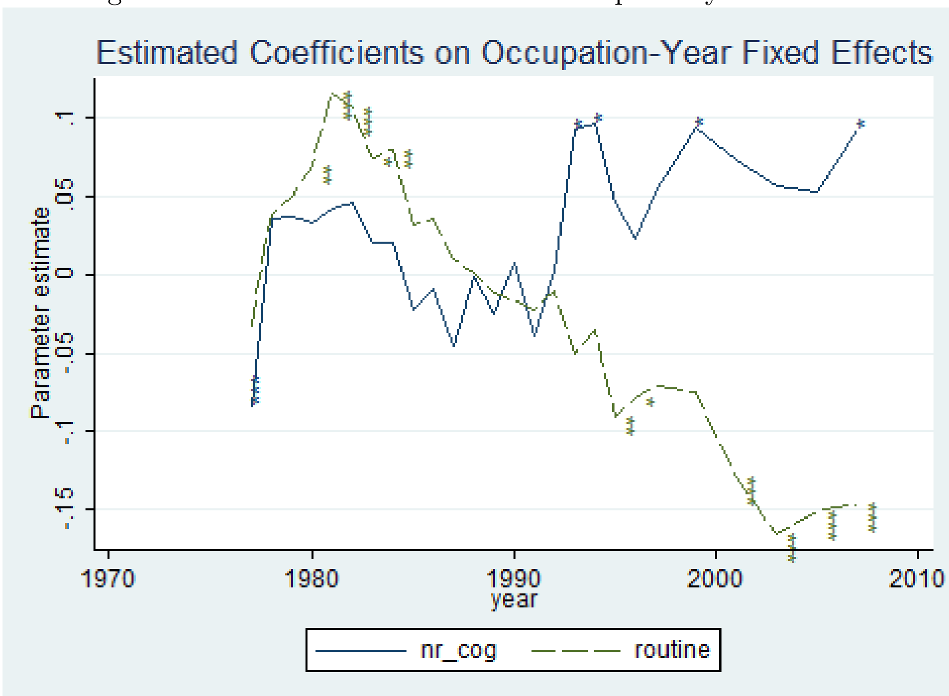
[48]The standard errors in Table 2.9 are obtained through the same bootstrap procedure as those in Table 2.2, except that now the first step that is bootstrapped is the estimation of Equation (2.34) instead of Equation (2.28), and ability rankings are built based on $X_i\widehat{\beta}_t + \widehat{\nu}_{ij}$. See footnote 38 for more details on the bootstrap procedure.

Table 2.9: Regressions of the probability of switching two years ahead for routine workers (odd years only, 1977-2005)

	P(sw)	P(nr_cog)	P(nr_man)
	(1)	(2)	(3)
q-1	.033	-.0009	.034
	(.012)***	(.010)	(.006)***
q-2	-.008	-.011	.003
	(.011)	(.010)	(.004)
q-4	-.001	-.004	.003
	(.011)	(.010)	(.004)
q-5	.091	.102	-.011
	(.013)***	(.013)***	(.003)***
Const.	.122	.102	.020
	(.008)***	(.007)***	(.003)***
Obs.	12530	12530	12530

Note: Sample includes workers in routine occupations in year t. In Column (1) the dependent variable is a dummy equal to 1 if the worker is employed in any non-routine occupation in year $t+2$. In Column (2) the dependent variable is a dummy equal to 1 if the worker is employed in a non-routine cognitive occupation in year $t+2$. In Column (3) the dependent variable is a dummy equal to 1 if the worker is employed in a non-routine manual occupation in year $t+2$. q-1 through q-5 represent dummies for the individual workers' estimated ability quintiles among routine workers in year t (with q-1 representing the lowest ability workers and q-5 the highest), obtained from the estimation of Equation (2.34). Workers in the middle of the ability distribution (q-3) are the omitted category. Standard errors are bootstrapped using 100 replications based on 6975 clusters of individuals. See footnotes 38 and 48 for details.

Figure 2.9: Estimated coefficients on occupation-year dummies



Note: The figure shows the estimates of $\widehat{\theta}_{rt}$ and $\widehat{\theta}_{cog}$ obtained from the wage equation (2.34). Stars denote the level at which the estimated coefficients are significantly different from zero.

return to non-routine cognitive jobs that was observed in Figure 2.6. However, the pattern remains such that the wage premium in routine occupations experiences a substantial fall relative to the wage premium in either of the non-routine occupational categories. Finally, Table 2.10 confirms that the results in terms of wage changes for switchers presented for the baseline specification also go through when allowing for changing returns to education.

2.7.3 Alternative Classification of Occupations

Appendix A.3 describes an alternative classification of occupations into the broad categories (non-routine manual, routine, non-routine cognitive) based directly on task information from the Dictionary of Occupational Titles, rather than a simple aggregation of 3-digit codes. The alternative classification procedure is related to Autor et al. (2003) and Autor and Dorn (2009),

Table 2.10: Wage changes for routine workers, according to direction of switch

Panel A: Dependent variable is change in fitted model wages (in logs)

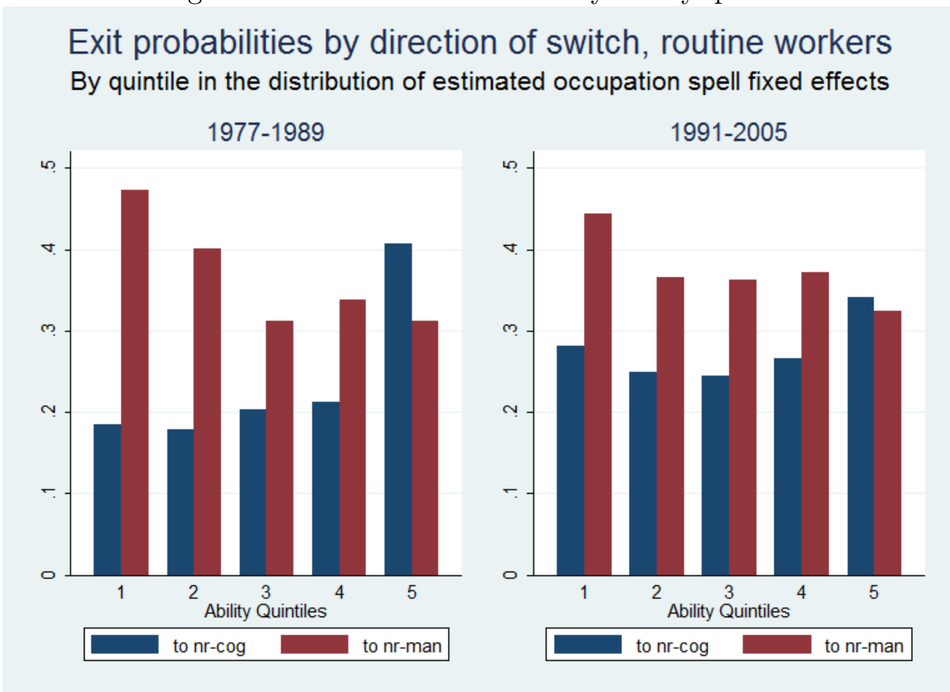
	Change in fitted model wages between year t and year:					
	$t+1$	$t+2$	$t+4$	$t+10$	$t+2$	$t+2$
Period:	1976-1997	1976-2007	1976-2007	1976-2007	1976-1989	1990-2005
	(1)	(2)	(3)	(4)	(5)	(6)
goto-nrcog	.089	.124	.100	.144	.069	.170
	(.010)***	(.009)***	(.008)***	(.011)***	(.012)***	(.011)***
goto-nrman	-.147	-.135	-.027	.062	-.146	-.126
	(.023)***	(.021)***	(.020)	(.028)**	(.032)***	(.027)***
Const.	-.042	.032	.068	.018	.034	-.023
	(.002)***	(.003)***	(.004)***	(.009)**	(.003)***	(.004)***
Obs.	15749	18188	14179	7548	9358	8830
# of Indiv.	2632	3181	2649	1722	1808	2355
R^2	.171	.178	.15	.091	.184	.198

Panel B: Fraction of routine workers in each of the switching categories (%)

	Fraction of routine workers in year t switching to non-routine jobs in year:					
	$t+1$	$t+2$	$t+2$	$t+2$	$t+2$	$t+2$
Period:	1976-1997	1976-2007	1976-2007	1976-2007	1976-1989	1990-2005
	(1)	(2)	(3)	(4)	(5)	(6)
goto-nrcog	8.05	10.83	11.15	11.43	9.44	12.31
goto-nrman	1.52	2.18	1.93	1.87	1.83	2.56

Note: Workers who stay in routine occupations are the omitted category. All regressions include year dummies. The wage changes are taken over the time horizons indicated above each column (in years). For column 1, occupation transitions between years t and t+1 are considered. For column 2 onwards, occupation transitions between years t and t+2 are considered (even though the wage change may be taken over a longer horizon). Standard errors are clustered at the individual level.

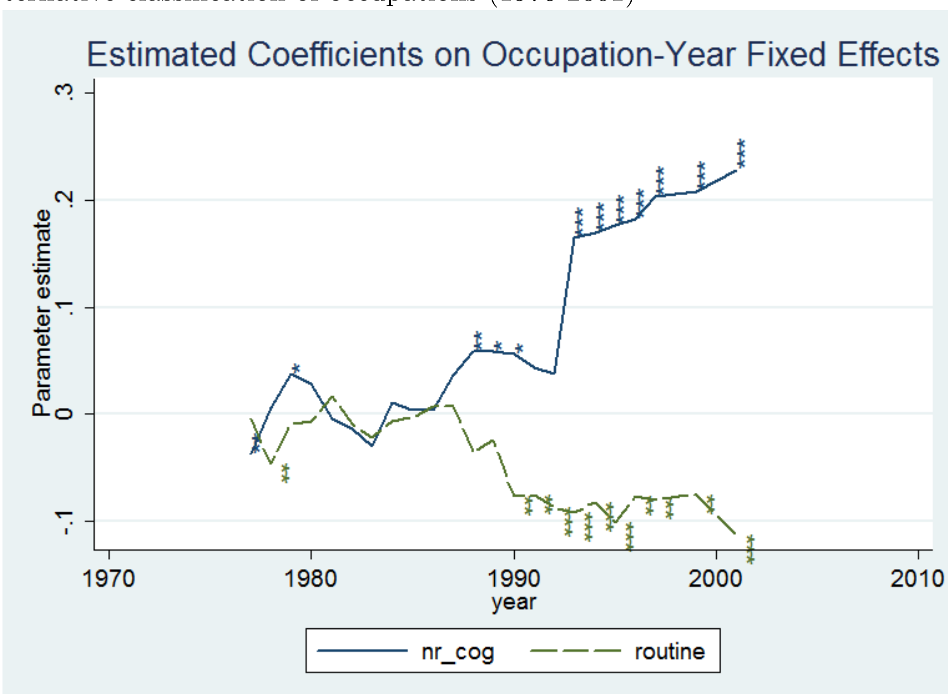
Figure 2.10: Direction of switch by ability quintile



Note: Sample includes workers in routine occupations, and plots their probability of switching to the different non-routine occupations between years t and $t+2$, according to their ability quintile, using the alternative classification of occupations described in Appendix A.3.

and is applicable to the data up to 2001. Figure 2.10 plots the switching probabilities for routine workers into the two non-routine categories. With this alternative classification of occupations the measured switching rates are higher, but the general pattern in the direction of switch across ability quintiles remains robust. Meanwhile, Figure 2.11 shows the estimated changes in occupation wage premia from the estimation of Equation (2.28) using the alternative occupation classification. The Figure confirms the robustness of the result on the falling wage premium in routine occupations relative to either non-routine category.

Figure 2.11: Estimated coefficients on occupation-year dummies under alternative classification of occupations (1976-2001)



Note: The figure shows the estimates of $\widehat{\theta}_{rt}$ and $\widehat{\theta}_{cog}$ obtained from the wage equation (2.28) using the alternative classification of occupations into broad categories described in Appendix A.3. Stars denote the level at which the estimated coefficients are significantly different from zero.

2.8 Conclusions

This paper derives the individual-level effects of routinization-biased technical change in a model of occupational sorting, and provides empirical evidence on the labor market experience of workers holding routine jobs during the past three decades in the United States.

Consistent with the prediction of the model, the data show strong evidence of selection on ability in occupational mobility out of routine occupations: workers of relatively high ability are more likely to switch to non-routine cognitive jobs, and workers of relatively low ability are more likely to switch to non-routine manual ones. Interestingly, after the 1990s, the probability of switching to non-routine cognitive jobs increases more than the probability of switching to non-routine manual jobs for routine workers at all ability quintiles. This suggests that there has not been a large displacement of middle-skill workers towards low-skill jobs in the 1990s or 2000s, as has been sometimes assumed.

In terms of wage growth, also consistent with the prediction of the model, workers staying in routine jobs perform significantly worse than workers staying in any other type of occupation. This is due to a substantial fall in the wage premium for routine occupations, which is not driven by changes in the composition of the workers in terms of their skill level, or by changes in the return to education. Meanwhile, switchers from routine to non-routine manual suffer substantial wage cuts relative to stayers over horizons of up to two years, but those wage cuts disappear over time. Workers who switch from routine to non-routine cognitive jobs have significantly higher wage growth than stayers over a variety of time horizons.

The results in the paper provide micro-level evidence for the dynamics underlying the aggregate patterns of changes in employment shares and mean wages across major occupation groups. The paper finds that a model of occupational sorting with routine-biased technical change is able to rationalize many of the individual-level facts concerning the labor market experience of routine workers in the United States in the past three decades. The fact that there is a continuum of skills and three occupations, as in Jung and Mercenier (2010), rather than a continuum of occupations and three skill groups, as in Acemoglu and Autor (2011), is crucial, as it allows for the (empirically relevant) differentials in the wage changes for workers with different occupational switching trajectories.

Thinking of occupational mobility and wage changes in the context of routine-biased technical change gives a new perspective on the findings in the literature on occupational mobility which uses individual-level data. For example, part of the increase in mobility found in Kambourov and Manovskii (2008), and part of the U-shaped pattern in occupational switching found in Groes et al. (2009) may be related to the process of routinization.

In future work, it would be interesting to formally decompose the changing aggregate shares of employment of the three major occupation groups into changes in exit rates and changes in the patterns of occupational choice of new entrants to the labor market, in order to have a more precise estimate of the extent of job displacement that could be attributed to routine-biased technical change.

Another important avenue for future work would be to consider the implications of occupation-specific skills (Kambourov and Manovskii, 2009b) within the context of the model discussed in this paper. Workers that have built more human capital specific to routine jobs (generally older workers) will have a lower incentive to switch out of those jobs, even if the prospective wage growth in their occupation is low, as they would face a large human capital loss from switching occupations. This consideration would be less important for younger workers. The ability cutoffs at which young and old routine workers decide to switch occupations, then, would not be the same.

Extending the model in this way would induce a dynamic aspect to workers' switching decisions. A worker with specific human capital who switches occupations would be trading off an initial loss of specific human capital against the access to an occupation with a steeper wage profile. This type of consideration could explain the findings in this paper regarding the wage changes for switchers to non-routine manual occupations. They experience an initial wage cut but subsequently recover from it, as wages in non-routine manual occupations are growing faster than they do in routine ones.

Chapter 3

Task Distance: Construction, Trends and Relationship with Wage Changes for Individual Workers[49]

3.1 Introduction

Understanding how individuals build their human capital has been an issue of great interest in economics. The extent to which an individual's human capital is specific (to a firm or an industry, for example) or general (and therefore transferable across jobs) has important implications for the analysis of an economy's reaction to shocks and the definition of appropriate policy responses to issues such as unemployment.

Given the positive correlation between employer tenure and wages, a traditional approach has been to consider human capital as having an important employer-specific component.[50] Neal (1995) and Parent (2000), however, have suggested that human capital should be characterized as industry-specific: firm tenure is irrelevant once industry tenure is taken into account. Kambourov and Manovskii (2009b), meanwhile, have recently presented evidence that suggests that human capital is actually occupation specific. Using data from the PSID, they find substantial returns to occupational tenure, and find that industry and employer tenure have relatively little effect on wages once occupational tenure is controlled for. Interestingly, in Kambourov and Manovskii (2008), they also find that the U.S. labor market

[49]This chapter benefited from multiple discussions with Giovanni Gallipoli, Nicole Fortin, David Green and Thomas Lemieux, as well as with participants at the UBC Empirical Lunch and the CEA Conference (2010). Financial support from CLSRN is gratefully acknowledged.

[50]See, for example, the evidence presented in Topel (1991) for the U.S., or in Dustmann and Meghir (2005) for West Germany.

experienced an important increase in the rate of worker mobility across occupations over the period 1968-1997.

Another set of papers have suggested a more nuanced view of human capital transferability. They argue that the extent to which human capital is transferable across jobs should depend on the set of skills that the worker possesses, and the set of tasks that the jobs require. An example of this strand of the literature is Lazear (2003). In his view, there exist a set of worker skills that are used in many firms, but each firm differs in the extent to which it uses each particular skill. Gathmann and Schonberg (2010) meanwhile develop the concept of "task-specific human capital". In their model, skills accumulated in an occupation are only productive in occupations in which a similar combination of tasks is performed. When a worker switches occupations, he loses his occupation-specific human capital, but is able to transfer a fraction of his task-specific human capital. The fraction that can be transfered will be proportional to the degree of similarity between the two occupations in terms of their task content. They use data from Germany to generate distance measures across occupations based on the (dis)similarity of the mix of tasks performed. Combining this information with a panel of administrative data for German employees, they find that task-specific human capital accounts for an important fraction of wage growth. They also find evidence that workers do not switch occupations randomly, as they are disproportionately more likely to go to occupations with similar task requirements.

Poletaev and Robinson (2008) have also characterized occupational switching according to the implied changes in task content. They follow Ingram and Neumann (2006) in estimating a factor model to determine the "basic skills" that define an occupation. Their data source is the Dictionary of Occupational Titles (DOT) of the United States. They define a discrete measure of skill switching and apply it to the set of occupational switchers observed in the Displaced Worker Survey of the U.S. The authors find that the wage cut experienced by 'skill switchers' is significantly larger than for 'skill stayers'. Moreover, they find that, among skill switchers, the wage cut does not vary between industry switchers and industry stayers. However, among skill stayers, industry switchers experience substantially larger wage cuts than industry stayers. This evidence seems to suggest that workers build industry-specific human capital, but only within their skill set. When the set of skills required for the job changes, their industry experience becomes irrelevant.

In this paper, I characterize occupations by a set of skill requirements, and follow Gathmann and Schonberg (2010) in developing a distance measure between occupations. This allows for a more detailed characterization of occupational mobility, by distinguishing between moves among similar occupations and moves among more different occupations in terms of their skill requirements.

I analyze the evolution over time in the distance of observed occupational transitions. In an environment of rising occupational mobility, I address the question of whether the increase in occupational mobility (documented in Kambourov and Manovskii (2008)) has been driven by transitions between more similar or more different jobs.[51] To do this, I merge data on the skill content of occupations from the Dictionary of Occupational Titles (DOT) with data on observed occupational transitions in the Panel Study of Income Dynamics (PSID). A major advantage of these datasets as compared to those used by Gathmann and Schonberg (2010) is that the occupational classifications have a much finer level of detail. Their dataset has only 64 different occupational codes. The data used here, which uses 3-digit Census Occupation Classification (COC) codes, has a universe of over 400 occupations. This dataset therefore captures a much larger number of switches between detailed occupations. The finer level of occupational detail also implies that the task measures should provide a more precise characterization of the task content of each particular occupation.[52] [53]

I find that there is a small but statistically significant increase in the distance of occupational switches between 1968 and 1997. This means that in the 1990s, not only are people switching occupations more frequently, but they are also switching to occupations that are (slightly) more different in terms of their ability requirements and their task components.[54] I find that

[51] See Robinson (2011) for a recent paper with a similar goal.

[52] To get task measures at the level of their occupational codes, Gathmann and Schonberg (2010) calculate the fraction of people in each occupation who perform a given task (using individual-level data on task performance). These fractions are rarely zero or one, which implies that a single occupational category is combining a number of workers performing different mixes of tasks, presumably because they work in a number of different specific (more disaggregated) occupations. (Robinson (2010) expresses similar concerns.) In my dataset, on the other hand, I do not have information on the fraction of people performing a given task in each specific occupation. Rather, I have information on the importance of different tasks for the appropriate performance of an occupation, and I assume all workers in the occupations perform these tasks at that level. Since the set of occupations I use is much more detailed, I expect these measures to be fairly accurate.

[53] In addition to Ingram and Neumann (2006) and Poletaev and Robinson (2008), other papers characterizing occupations according to their task content using the DOT or its successor, O*Net, include Autor, Levy, and Murnane (2003), Peri and Sparber (2009), and Firpo, Fortin, and Lemieux (2011), among many others.

[54] In the Appendix I discuss some of the challenges that arise when trying to perform a

occupational switches exhibit significantly larger distances for more educated workers, white workers, females, and workers living in larger cities. The positive trend over time in distance remains significant even after controlling for a number of demographic characteristics. Therefore, the result is not driven by composition effects.

I also analyze the impact of the distance of occupational switches on wages. If human capital is specific to a set of skills, one would expect workers to experience larger wage losses if they experience a larger change in the mix of tasks they perform. I do not, however, find a robust direct negative effect of distance on wages. Instead, I find some evidence that wage changes are negatively correlated with distance only for particular subsets of workers, namely workers who experience occupational downgrades, defined as decreases in their intensity of abstract non-routine task usage. This result suggests that a measure of task distance may provide a good metric of the costs of switching occupations for certain workers (such as displaced workers, presumably), but not for others, such as workers experiencing promotions or moving to jobs that better match their skill set. Caution should therefore be used when applying a distance measure to a broad range of occupational transitions as it may not provide an accurate general metric of transition costs.

The rest of the paper is organized as follows: Section 3.2 describes the datasets and methodologies for developing distance measures and merging them to observed occupational transitions. Section 3.3 documents the evolution over time of the distance of occupational switches in the U.S. and analyzes cross-sectional differences in distance. Section 3.4 presents evidence on the individual-level relationship between distance and wage changes. Section 3.5 concludes and suggests ideas for future work.

3.2 Data and Methodology

This sections describes the data sources and the procedures used to match occupational descriptions with data on observed occupational transitions.

similar analysis using data from the Current Population Survey (CPS). (See also Robinson (2010)). The increasing trend in distance seems not to be present in the CPS. It remains as future work to verify why the CPS does not yield similar results to the PSID.

3.2.1 PSID Data

Description

The Panel Study of Income Dynamics (PSID) is a national longitudinal study of nearly 9,000 U.S. families. Following the same families since 1968, the PSID collects data on economic, health, and social behavior.[55] The main advantage of the PSID is its wide longitudinal dimension, tracking the same set of individuals every year between 1968 and 1997.[56] The PSID data also permit the construction of consistent series of occupational affiliation for the whole period, using the 1970 Census classification (COC).[57] The main disadvantage of the PSID is its relatively small sample size (particularly when considering particular age-education groups or particular occupational categories).[58]

Occupational transitions

The PSID records occupation data for household heads and wives. Occupational transitions can be computed in the PSID by taking advantage of the survey's longitudinal dimension, comparing an individual's current occupation report with his most recent report in the previous surveys.[59]

As described in Kambourov and Manovskii (2008), the PSID did not originally code occupations and industries at the three-digit level prior to

[55] PSID data is publicly available at `http://psidonline.isr.umich.edu/` The Panel Study of Income Dynamics is primarily sponsored by the National Science Foundation, the National Institute of Aging, and the National Institute of Child Health and Human Development and is conducted by the University of Michigan.

[56] Starting in 1997 the PSID changed its methodologies and began to interview individuals every two years.

[57] The PSID classifies occupations using 1970-COC up to 2001, and using 2000-COC from 2003 onwards. Note that a downside of the PSID's method of consistently using the 1970-COC up until 2001 is that this classification may not be very appropriate for latter years. Specifically, workers in new occupations that did not exist in 1970 (and are therefore not present in the 1970 classification scheme) would be coded into different "not elsewhere classified" categories. Kambourov and Manovskii (2008) document that the fraction of workers employed in the "not elsewhere classified" categories in their PSID sample increases from 14% to 21% over the 1968-1997 period.

[58] The Current Population Survey (CPS) provides a much larger sample size, but may be problematic for an analysis of occupational mobility. This is discussed further in Appendix B.2.

[59] In this paper, I consider switches by comparing occupation reports in two consecutive years. However, it would also be possible to construct switches over longer horizons, in order to account for occupational transitions that involve intervening periods of unemployment (something that is not possible in the CPS, where data is available on the current and previous occupation only for people who were working in both periods).

1981. In 1999, they released the "1968-1980 Retrospective Occupation-Industry Supplemental Data Files", in which they recoded these variables at the three-digit level for those survey years. To produce the three-digit recode, the PSID pulled out paper materials from its archives containing the written records of the respondents' descriptions of their occupations and industries. Using these records, the PSID assigned three-digit 1970 Census codes to the reported occupations and industries of household heads and wives for the period 1968-1980. To save time and increase reliability, the coder coded all occupations and industries for each person across all required years before moving on to the next case. Thus, in constructing the Retrospective Files, the coders had access not only to the respondents' description of their current occupation (industry) but also to the description of their past and future occupations (industries). This allowed them to compare these descriptions, decide whether they were similar, and assign the same occupation (industry) code where appropriate. This coding methodology minimizes coding error and provides a reliable series of occupation codes for household heads and wives in the PSID.[60]

From 1981 onwards, unfortunately, only the originally coded data is available in the PSID. This period is therefore plagued with coding error, stemming from the fact that the coder did not have access to the individual's description of their occupation in the previous or subsequent years in order to make a comparison and determine whether a switch had in fact occurred. This results in a much higher number of occupational transitions being recorded relative to the period before 1981. To deal with the break in the series that occurs due to the change in the coding procedure, I analyze the periods before and after the change separately, that is, I consider the periods 1968-1980, and 1981-1997.

3.2.2 DOT Data

Description

The Dictionary of Occupational Titles characterizes occupations along a series of common dimensions. The 4th Edition of the DOT was published in 1977 and is available in electronic format through the Interuniversity Consortium for Political and Social Research (ICPSR, 1981). A Revised 4th Edition was published in 1991 (ICPSR, 1991). The DOT dataset provides precise measures of the different abilities that are required in different

[60]See the documentation on the PSID's website, as well as Kambourov and Manovskii (2008) for more details on this argument.

occupations, as well as the different work activities performed by job incumbents. In this paper, this dataset will be used in order to build empirical measures of task distance between occupations, by quantifying the degree of dissimilarity across the task vectors of different occupation pairs.

DOT Dimensions

The dimensions along which the DOT dataset characterizes occupations are the following (ICPSR, 1981):

- 3 Complexity of work variables (in relation to Data, People, and Things).
- 3 General Educational Development (GED) variables (Reasoning, Mathematical and Language)
- 1 Specific Vocational Preparation measure (amount of time required to learn tasks)
- 11 Aptitudes
- 10 Temperaments
- 6 Physical demands
- 7 Environmental conditions
- 5 Interests

Table 3.1 provides an example of the ratings along the GED and Aptitude dimensions for 5 different occupations (see below for details on score normalization).

3.2.3 Matching DOT with PSID

The DOT-77 and DOT-91 each have their own coding schemes, which are much more disaggregated than the COC. To aggregate to the 1970-COC level, I follow Autor et al. (2003). I use the April 1971 CPS Monthly File, in which experts assigned individuals both with 1970-COC and DOT-77 codes. Using the CPS sampling weights, I calculate means of each DOT task measure at the 1970-COC occupation level. For DOT-91, I first convert the codes to DOT-77 codes using a crosswalk provided by David Autor. I then follow the same procedure as with DOT-77 in order to obtain DOT-91 scores at the 1970-COC level. Each DOT score is rescaled to have a (potential) range from zero to 10.

Table 3.1: Examples of DOT-77 scores (rescaled to 0-10 range)

Occupation	Economists	Element. school teachers	Sales clerks, retail	Auto mechanics	Truck drivers
Reasoning	7.81	7.97	5.04	5.78	3.93
Math	6.59	4.09	3.85	3.83	1.56
Language	7.40	7.92	4.17	3.91	2.03
Intelligence	7.42	7.49	5.02	5.00	4.89
Verbal	7.36	7.49	4.98	4.91	2.60
Numerical	6.83	4.97	4.91	2.84	2.58
Spatial	3.50	2.54	2.94	6.99	5.19
Form Percep	3.63	4.92	3.36	4.97	2.84
Clerical Per	6.19	7.33	4.41	2.61	2.34
Motor Coord	2.66	2.67	4.71	4.86	4.91
Finger Dext	2.70	2.65	3.85	4.89	2.56
Manual Dext	2.54	2.63	3.85	7.13	4.93
Eye-Hand-Foot	0.08	2.28	0.06	2.31	4.86
Color Discrim	0.45	4.46	2.69	2.28	2.39

3.2.4 Selection of Occupational Characteristics

The choice of the relevant dimensions with which to characterize occupations is somewhat arbitrary. I choose the three GED variables and the eleven aptitudes as the relevant dimensions. In the Appendix, I present results using different subsets of data from O*Net, the successor to the Dictionary of Occupational Titles. The findings are robust to these different specifications.

3.2.5 Distance Measures

My goal in developing a measure of distance across occupation pairs is to capture how different the two occupations are in terms of their task content and their ability requirements. The measure I consider is an Euclidean distance. Let x_o^a be the importance score of dimension a (which will be each of the dimensions mentioned in Section 3.2.4) in occupation o. $x_{o'}^a$ is the analogous measure for occupation o'. The distance is measured as:

$$D_{o,o'} = \sqrt{\sum_{a=1}^{A} \left(x_o^a - x_{o'}^a\right)^2} \tag{3.1}$$

Table 3.2: Examples of some observed low- and high-distance occupational transitions

	Low distance		High distance	
	Cashiers	Retail Salesp.	Secretary	Waiter
DOT-77:				
Distance	4.09		11.68	
Reasoning	4.23	5.04	6.00	3.84
Math	2.89	3.85	3.98	1.99
Language	2.25	4.17	5.96	2.47
Intelligence	5.21	5.02	7.42	4.71
Verbal	5.16	4.98	7.42	4.70
Numerical	5.19	4.91	5.00	2.59
Spatial	2.45	2.94	2.51	2.50
Form Percep	4.02	3.36	7.30	2.57
Clerical Per	5.23	4.41	7.41	2.54
Motor Coord	6.29	4.71	7.32	3.20
Finger Dext	6.25	3.85	7.32	2.61
Manual Dext	4.76	3.85	4.92	4.89
Eye-Hand-Foot	0.03	0.06	0.02	2.30
Color Discrim	1.76	2.69	2.39	0.24

where A is the total number of dimensions being considered.

In some cases, I work with what I call "log distance", which I define as:

$$d_{o,o'} = \ln(1 + D_{o,o'}) \tag{3.2}$$

Table 3.2 presents an example of two occupational transitions: a transition between Cashier and Retail Salesperson - which is characterized as being low-distance - and a transition between Secretary and Waiter - which is characterized as high distance. The Table also shows the scores for the DOT characteristics used to calculate the distance measure.

3.3 Findings: Trends and Cross-Sectional Variation in Task Distance

In order to study the variation in the the task distance of occupational transitions for different demographic groups, as well as its changes over time, I attach the measure of distance described above to observed occupational transitions in the PSID. I restrict the PSID sample to household heads and wives between 19 and 70 years old.[61] For this sample, there are a total of 61,107 occupational transitions over the 1968-1997 period. The transitions are not distributed evenly over time: there are many more transitions from 1981 onwards, due to the effect of coding error (see Section 3.2).

Figure 3.1 shows the median annual distance in the sample of occupational switchers for the measure of distance based on all GED and Aptitude dimensions of the DOT-77 and DOT-91. The data for 1981 onwards are slightly lower in level and are therefore graphed on the second y-axis. This is due to the coding error that affects this period. The coding error reduces the measured level of distance because, when a worker has a switch that is due to coding error (i.e. the worker stayed in the same occupation, but the description of the occupation receives a different code in two consecutive years), the incorrect code is likely to reflect an occupation that is similar (according to the distance measure) to the true (pre-switch) occupation. Thus, switches due to coding error tend to be switches with a small distance. Since I am measuring the distance conditional on switching, the addition of a large number of low-distance switches due to coding error from 1981 onwards reduces the measured median distance for that period. The Figure shows an increasing trend in median distance both before and after 1981. A simple regression of the distance measure on a time trend yields a positive and significant coefficient for each of the sub-periods.

To test the changes over time more formally, I run a regression of distance on a number of demographic characteristics, as well as a time trend. This serves two purposes. First, the correlates of distance are interesting in their own right, in order to establish which worker characteristics are associated with larger or smaller distances. Second, I am interested in ruling out the possibility that the increase in distance over time is due to a composition effect. That is, I am interested in testing whether the derivative of distance with respect to time is still significantly positive once I control for the composition of the workforce along several observable characteristics.

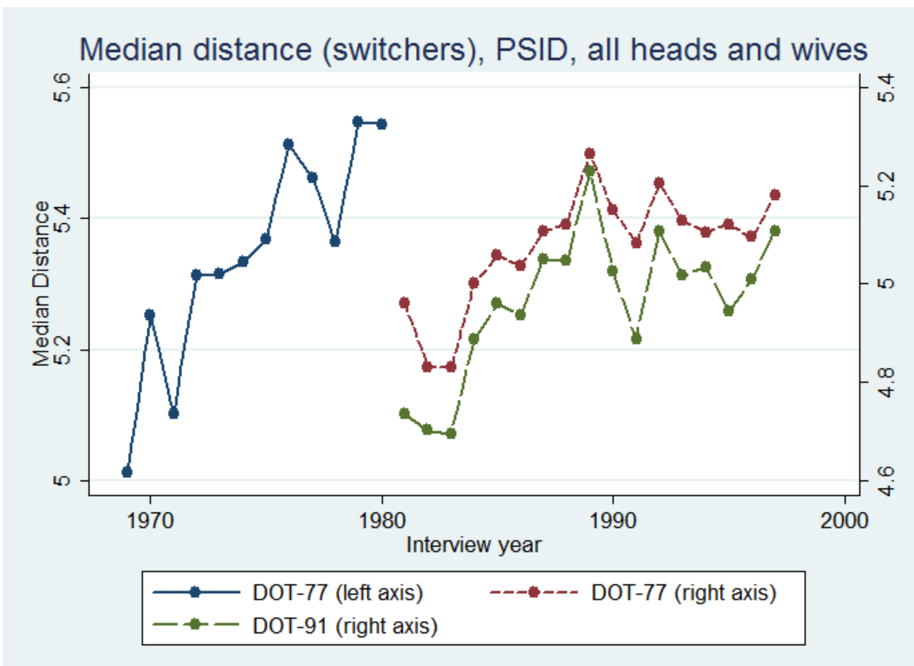
I run an OLS regression of the following form:

$$d_{it} = \beta_0 + \beta_1 t + \beta_2 X_{it} + \epsilon_{it} \tag{3.3}$$

where d_{it} is log-distance between the worker's occupation in the previous year and the worker's occupation in year t, where the distance measure is

[61] I keep individuals who are assigned a weight of zero by the PSID, such as the Latino sample added in the 1990s. Excluding zero-weight observations does not significantly alter the results.

Figure 3.1: Median distance for occupational switchers, all household heads and wives



constructed using DOT-77 as described above. t is a linear time trend, and X_{it} includes demographic controls, namely dummies for females, non-whites, individuals with at least some college education, married workers, workers living in metropolitan areas (SMSA), as well as age and age squared. I estimate the equation using only occupational switchers (and therefore the dependent variable should be thought of as the distance conditional on switching occupations). Because coding error may affect different demographic groups in different ways, rather than estimating the equation on a pooled sample for all years and including a dummy for the period from 1981 onwards, I instead estimate the equation separately for the 1969-1980 and the 1981-1997 periods.

The results are presented in Table 3.3. Column (1) is for the early part of the sample, 1969-1980. Columns (2) and (3) are for the latter part of the sample, from 1981 onwards. Column (3) addresses the coding error issues in the latter part of the sample by using only a restricted sample of switchers: those who also report an employer switch along with the occupational switch. These are more likely to be true occupational switchers.[62]

The results in the Table shed light on cross-sectional differences in the distance of occupational switches. In particular, significantly higher distances are observed among more educated workers, white workers, females, and workers living in larger cities. In terms of the time trend, there is evidence of a significant increase over time. Therefore, I can conclude that composition effects are not driving the result; there is a true unexplained increase in distance over the 1968-1997 period. It must be noted, however, that the magnitude is very small. The mean of log-distance is 1.83 for the sample in Column (1), 1.78 for the sample in Column (2), and 1.87 for the sample in Column (3), with standard deviations of 0.40, 0.44, and 0.39, respectively. Therefore, the coefficients on the time trend imply increases in distance of less than 0.008 standard deviations per year. In Appendix B.1 I present an extensive robustness check using O*Net data and an alternative methodology to address the coding error in the post-1981 period. That Appendix also plots mean distances over time and over the life cycle for different education groups.

The findings on the cross-sectional heterogeneity should be seen as motivating further research in order to understand whether the demographic groups that switch higher distances are doing so because they are able to transfer a larger fraction of their human capital across distant occupations,

[62] See Kambourov and Manovskii (2008) for a justification of this argument, and Appendix B.3 for information on how employer switches are identified in the data.

Table 3.3: Regression of distance on observables - PSID data

	69-80	81-97	81-97 R
	(1)	(2)	(3)
Time	.003	.003	.002
	(.001)***	(.0005)***	(.0009)**
Age	-.004	-.008	-.002
	(.003)*	(.002)***	(.003)
Age^2	.00003	.00007	1.81e-06
	(.00003)	(.00002)***	(.00004)
Female	.019	.039	.056
	(.010)*	(.006)***	(.009)***
College	.090	.026	.030
	(.011)***	(.006)***	(.009)***
Non-White	-.097	-.030	-.038
	(.011)***	(.007)***	(.009)***
Married	.012	-.001	.009
	(.010)	(.006)	(.009)
SMSA	.048	.023	.021
	(.011)***	(.006)***	(.009)**
Obs.	10916	48025	10215
R^2	.041	.009	.011
e(N-clust)	4560	12854	6015

Note: Standard errors clustered at the individual level. e(N-clust) denotes the number of clusters (individuals). Dependent variable is log-distance (based on DOT-77 GEDs and Aptitudes). 'R' indicates restricted switches (occupation switches accompanied by employer switches).

or because they are suffering from adverse shocks that force them to accept high-distance switches that entail large human capital losses. The next section of the paper presents further motivating evidence by studying the relationship between task distance and the wage changes of individual workers.

3.4 Relationship Between Task Distance and the Wage Changes of Individual Workers

In this section I explore whether the wage change experienced by workers who switch occupations is correlated with the distance of their switch. The sample is further restricted to workers with valid wage reports in two consecutive years, with real hourly wages in 1979 dollars between \$1.1 and \$54.6 (in order to exclude outliers). I exclude people assigned a PSID sample weight of zero, as well as people who are ever self-employed, or ever in the military. Throughout this section I use "log distance" based on the DOT-77 as the distance measure.

I first confirm the results presented by Gathmann and Schonberg (2010) regarding the auto-correlation of wages. As in their paper, I regress current real hourly log wages on their lagged value.[63] I do this separately for occupation stayers and occupation switchers, and I also test the effect of distance on this auto-correlation. The results are presented in Table 3.4. Column (1) includes workers who do not switch occupations. Columns (2) and (3) are for occupation switchers over the period 1971-1980. Columns (4) and (5) are a similar sample for 1981-1997. As this latter period is plagued with coding error, I also show results in Columns (6) and (7) when I restrict the occupational switchers to those who also switch employers.

Column (1) shows that wages are highly auto-correlated for workers who do not switch occupations. The auto-correlation is significantly lower for occupational switchers, as is clear from columns (2) and (6). Moreover, the auto-correlation of wages falls with distance, as evidenced by the negative and significant coefficient on the interaction of lagged wages and distance in Columns (3), (5) and (7). In fact, for low-distance occupational switchers, the auto-correlation is similar to that of occupational stayers. Table 3.5 presents the estimated auto-correlation of wages, calculated at different percentiles of the log-distance distribution.

In Table 3.6 I use the regression results to estimate the marginal effect of distance on current wages at different percentiles of the distribution of lagged wages. The values can be interpreted as elasticities. The estimated marginal effect is positive for people on the low end of the lagged-wage distribution, while it is negative and significant for workers on the top half of the distribution.

In the rest of this section, I explore the relationship between distance and wage changes further. First, let the variable task intensity be defined

[63] See Appendix B.3 for details on the wage variable.

Table 3.4: Autocorrelation of wages and relation to distance - Dependent variable: Log current hourly real wage

	71-97	71-80		81-97		81-97 R	
	Stay	Switch	Switch	Switch	Switch	Switch	Switch
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Prev W	.875	.688	.860	.811	.908	.554	.910
	(.005)***	(.023)***	(.068)***	(.008)***	(.019)***	(.021)***	(.071)***
Prev W · Dist			-.097		-.057		-.199
			(.038)**		(.011)***		(.038)***
Dist			.143		.070		.201
			(.063)**		(.019)***		(.050)***
Obs.	34335	2265	2261	16199	16096	3127	3107
R^2	.881	.667	.669	.805	.806	.588	.597
e(N-clust)	5592	1263	1263	4212	4202	1911	1899

Note: Prev W is log real wage in year $t-1$. Dist is log distance. All regressions also include age, age^2, dummies for gender, education, race, marital status and metropolitan area residence, as well as year dummies and 2-digit occupation dummies (for destination occupation). Standard errors clustered at the individual level. e(N-clust) is the number of clusters (individuals). 'R' indicates restricted switches (occupation switches accompanied by employer switches).

Table 3.5: Estimated auto-correlation of wages at different percentiles of distance

Percentile	Auto-correlation of wages	
of log-dist	1971-1980	1981-1997 R
Stayers	.875	
1	.791	.747
10	.738	.644
25	.712	.587
50	.685	.536
75	.661	.485
90	.640	.449
99	.609	.387

Note: Auto-correlations estimated using the results presented in Table 3.4. Percentiles of log-distance are calculated separately for the relevant subsample in each column. All the estimated auto-correlations are significantly different from zero at the 1% level.

Table 3.6: Estimated marginal effect of distance on current wages at different percentiles of lagged wages

Percentile	Marginal Effect of Distance	
of Prev W	1971-1980	1981-1997 R
1	.069**	.085***
10	.038	.033
25	.019	.001
50	-.008	-.053***
75	-.038**	-.120***
90	-.059***	-.187***
99	-.093***	-.319***

Note: Marginal effects estimated using the results presented in Table 3.4. Percentiles of log-real wages in the previous year are calculated separately for the relevant subsample in each column. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

as follows:

$$TaskInt_o \equiv \sqrt{\sum_{a=1}^{A} (x_o^a)^2} \tag{3.4}$$

where x_o^a is the DOT score of dimension a in occupation o. In other words, I use a formula analogous to Equation 3.1 (the distance measure) to calculate the distance *from the origin* for each occupation. $TaskInt$ captures the overall intensity at which tasks are used in a particular occupation. The data shows a strong positive correlation between wages and the measure of task intensity.

When a worker switches occupations, he will experience changes in the level at which he performs tasks (captured by the change in $TaskInt$ between the worker's source and destination occupations), as well as changes in the specific mix of tasks he performs (captured by the distance measure, once the change in task intensity is controlled for). In the following regressions, I use the change in log real wages as the dependent variable (instead of the wage level). The first results are in Table 3.7. For the sample of occupational switchers, I regress the change in log wages on the distance measure, the log change in $TaskInt$, and a set of controls indicated at the bottom of the table. I run the regressions separately for the two sub-periods and present the results in Columns (1) and (2). As expected, I find that increases in

Table 3.7: Wage change regressions for occupational switchers

	Δw		$Abs(\Delta w)$	
	71-80	81-97 R	71-80	81-97 R
	(1)	(2)	(3)	(4)
Dist	-.010	-.049	.023	.058
	(.012)	(.015)***	(.011)**	(.012)***
$\Delta(TaskInt)$	.083	.094	-.017	-.012
	(.024)***	(.021)***	(.016)	(.016)
Obs.	2261	3107	2261	3107
R^2	.14	.209	.038	.038
e(N-clust)	1263	1899	1263	1899

Note: Wages changes are for hourly, real log-wages over two consecutive years. The sample includes only occupational switchers. Other controls are: lagged wages, age, age^2, and dummies for gender, education, race, marital status, metropolitan area residence, and year. Standard errors clustered at the individual level. e(N-clust) is the number of clusters (individuals). 'R' indicates restricted switches (occupation switches accompanied by employer switches).

task intensity are associated with wage increases. The effect of distance on the wage changes is negative, but insignificant for the 1971-1980 period.

In Columns (3) and (4), I instead use the absolute value of the wage change as the dependent variable. Here, I find that the effect of distance is positive and significant for both sub-periods. This means that higher distance switches are associated with larger wage changes (in absolute value).

Table 3.8 repeats the analysis adding a coarse indication of occupational downgrade. The variable *Down* is an indicator variable which is equal to one when a worker switches to an occupation with a lower use of abstract non-routine tasks. Following Autor et al. (2003), an occupation's use of abstract non-routine tasks is given by its average DOT importance score for "Direction, Control and Planning" and "GED-Mathematics". The Table shows that the negative effect of distance for the 1981-1997 period is only present among workers experiencing occupation downgrades, as defined here. The same is true for the positive effect of changes in task intensity in that period. Negative effects of distance are not found at all for the sample of workers who do not experience an occupational downgrade. To get an idea of the quantitative magnitudes, the coefficients from Column (4) in Table 3.8 imply that a one standard deviation increase in log distance for workers who experience an occupational downgrade (which is .389 for this subset of

Table 3.8: Wage change regressions

	1971-1980		1981-1997 R	
	Δw	Δw	Δw	Δw
	(1)	(2)	(3)	(4)
Dist	-.010	-.002	-.049	.018
	(.012)	(.020)	(.015)***	(.022)
$\Delta(TaskInt)$	.083	.061	.094	-.047
	(.024)***	(.042)	(.021)***	(.039)
Dist*Down		.0007		-.082
		(.029)		(.032)**
$\Delta(TaskInt)$*Down		.065		.171
		(.064)		(.059)***
Down		.006		.125
		(.048)		(.055)**
Obs.	2261	2261	3107	3107
R^2	.14	.141	.209	.213
e(N-clust)	1263	1263	1899	1899

Note: Other controls are: lagged wages, age, age^2, and dummies for gender, education, race, marital status, metropolitan area residence, and year. Standard errors clustered at the individual level. e(N-clust) is the number of clusters (individuals). Switchers in Columns (3) and (4) are restricted to those who are coded as switching occupations *and* employers. Down is an indicator variable equal to 1 for occupational transitions involving decreases in the use of abstract non-routine tasks.

workers) is associated with an additional wage fall of approximately 2.5% (all else equal including the change in task intensity).

Although the measure of occupational downgrade used here is very coarse, the results highlight the fact that there is not a robust negative effect of distance on wages. The human capital costs associated with an occupational switch that entails a larger distance might only be relevant for certain subsets of workers. In future work, it would be important to further study which workers are impacted by the distance of their switch by focusing, for example, on displaced versus non-displaced workers, or voluntary versus involuntary switches.

3.5 Conclusions and Future Work

This paper constructs a measure of distance between occupations based on information from the DOT, and merges it with data on observed occupational transitions from the PSID. I find that there is a slightly increasing trend in the distance of occupational switches between 1968 and 1997. This finding is not driven by composition effects. I also find some evidence of a negative correlation between distance and wage changes at the individual level, but only for a particular subset of workers.

The lack of a robust negative correlation between wage changes and distance points at the fact that there is substantial heterogeneity in the motives why workers switch occupations. Workers who make high distance switches may be doing so because they are able to transfer a larger fraction of their human capital across distant occupations, or it may be the case that they lose their job due to an adverse shock, and they accept a switch that involves a large loss of task-specific human capital because it dominates remaining unemployed. It may also be the case that certain workers benefit from a high-distance switch, because they experience an improvement in the quality of the match between their skills and the task requirements of the destination occupation, and these benefits outweigh the costs in terms of losses of task-specific human capital.

To obtain an unbiased estimate of the effects of distance on wages, assumptions need to be made about the underlying selection process, in order to account for the endogeneity of the decision to switch occupations. Gathmann and Schonberg (2010), for example, build on a Roy-type model of occupational sorting and develop an estimation strategy that explicitly accounts for selection. Another option is to focus on displaced workers, for whom the decision to switch occupations is more exogenous.

Overall, the results in the paper provide evidence on the importance of distinguishing between different types of occupational transitions. Theories of occupation-specific human capital that treat all occupation switches as equal may be ignoring substantial heterogeneity in the nature and implications of different types of switches.[64] Even a measure of task distance is limited in its ability to accurately capture the costs associated with a given switch, especially when selection is not carefully accounted for. More work is needed in order to fully understand the human capital implications of different types of occupational switches and their associated task changes.

[64]See also Robinson (2011) for evidence of important differences in the types of task changes observed for job switchers in the Displaced Worker Survey and in the March CPS.

Chapter 4

The Barriers to Occupational Mobility: An Aggregate Analysis[65]

4.1 Introduction

Several contributions to the human capital literature have analyzed the costs associated with different types of employment transitions. Topel (1991) provides evidence that a typical male worker in the United States with 10 years of job tenure would lose 25% of his wage if his job were to end exogenously. Dustmann and Meghir (2005), using data for West Germany, also find substantial returns to firm tenure, particularly for unskilled workers. Meanwhile, Neal (1995) and Parent (2000), suggest that workers lose a large fraction of their human capital when they switch industries, rather than firms.

Recently, Kambourov and Manovskii (2009b) have argued that human capital has a very important occupation-specific component. Support for this argument, at least among certain occupational groups, is also found in Sullivan (2010). Analyzing the costs of occupational mobility is particularly relevant in light of the evidence in Kambourov and Manovskii (2008) that occupational mobility in the U.S. is high and rose significantly between 1968 and 1997.

An adequate analysis of the costs of occupational mobility should take into account that occupational switches, which empirically are based on whether the occupation code assigned to a worker changes, may encompass many different types of transitions. In some cases a worker may be completely changing careers, while in others it may involve only a minor change

[65]This chapter benefited from discussions with Giovanni Gallipoli, Keith Head, Nicole Fortin and David Green, as well as with participants at the UBC Empirical Lunch, the UBC Macro Lunch and the LACEA Conference (2010). Carlos Sanchez provided valuable research assistance. Financial support from CLSRN is gratefully acknowledged.

in the set of tasks performed by the worker. Lazear (2003), Poletaev and Robinson (2008) and Gathmann and Schonberg (2010) argue in favor of linking human capital to tasks, and therefore differentiating occupational transitions according to the extent of task switching that they entail. Human capital built in an occupation should be partially transferable to other occupations where the set of tasks performed is similar.

The goal of this paper is to obtain an estimate of the costs of occupational mobility which takes into account the task content of occupations. The approach taken by the paper is somewhat unconventional relative to previous literature. I adapt a framework which is widely used in the trade literature in the form of 'gravity models'. In that literature, the interest is in estimating barriers to trade using data on flows of goods across countries, and proxies for trade costs that include geographical distance and whether the countries share a common border or a common language, among others. I show how the framework may be reinterpreted in order to consider the barriers to occupational mobility using data on the flows of workers across occupations, and proxies for mobility costs, mainly based on task data. Within the context of the model, the barriers to occupational mobility involve losses of human capital that are incurred when a worker needs to adapt the set of tasks that he is familiar with, or learn a new set of tasks altogether, upon switching occupations.

The theoretical framework provides a different approach to the analysis of the barriers to occupational mobility, and leads to the use of occupation-level rather than individual-level data. Previous models of occupational choice, such as Gibbons, Katz, Lemieux, and Parent (2005) or Groes, Kircher, and Manovskii (2009), have focused on the roles of comparative advantage across skill groups and learning in driving selection into occupations. Kambourov and Manovskii (2009a), meanwhile, have a model where occupational mobility is driven by persistent shocks to occupations that impact productivity of all workers in the occupation equally. In the model in this paper, the driving force behind occupational mobility is a set of non-persistent individual-specific productivity draws.

The framework in this paper involves a static, partial equilibrium model with perfect information. There is a continuum of homogeneous workers, who only differ in terms of the occupation that they start the period in (which is exogenously pre-determined). Workers make productivity draws from a set of extreme value distributions for each potential occupation. The extreme value assumption may be justified by thinking of workers as receiving a large set of offers from different employers within each occupation, and only considering the highest offer in each occupation. The distribution of 'highest offers' (maxima) across employers within each occupation would converge to an extreme value distribution.

Once workers observe their draws, they decide which occupation to work in during the period. There are costs to switching occupations, which depend on the particular occupation that the worker starts in, and the particular occupation that he considers switching to. This switching cost is not individual specific, but rather depends only on the identity of the worker's current and potential occupation. Based on his productivity draws, and taking into account the costs of mobility, the worker chooses the occupation where he will receive the highest wage. The workers' optimal switching decisions and the properties of the extreme value distribution lead to a 'gravity-type' equation, which predicts how the flow of workers across any given occupation pair is related to a set of occupation-specific characteristics, and to the cost of switching.

In order to estimate this gravity equation empirically, and to obtain an estimate of the barriers to occupational mobility, I need to define proxies for the costs of occupational mobility. I follow previous literature in characterizing occupations through a vector of skill or task characteristics (e.g. Autor, Levy, and Murnane (2003), Ingram and Neumann (2006), Gathmann and Schonberg (2010), Poletaev and Robinson (2008)). Using data from the Dictionary of Occupational Titles (DOT), I construct a distance measure between occupation pairs, which captures the degree of dissimilarity in the mix of tasks performed in the two occupations. If more specific human capital is lost when a worker experiences a more dramatic change in the set of tasks he performs, the barriers to occupational mobility should grow with distance. I also consider the role of two other proxies that might reflect task similarities not captured by the distance measure. The first is whether the two occupations share the same major occupational category, and the second is whether they share the same type of main task.

I then estimate the gravity equation empirically using data on worker flows across occupations from the Current Population Survey (CPS) for 1983 to 2002. Task distance is found to be a very important determinant of the cost of switching occupations, suggesting an important role for task-specific human capital. A conservative specification of the implied effects of task distance on the cost of switching suggests that a one standard deviation of distance increases the cost of switching occupations by 16.3% if the occupations are in different major occupational groups and do not share the same type of main task. The effect is reduced to 15% if the occupation pair share the same type of main task. If the occupations are in the same major occupation group, however, the implied cost of the transition is negligible under this specification.

I estimate several counterfactuals, including mobility rates out of particular occupations, and aggregate mobility rates. The estimates suggest that mobility rates would increase the least for workers in management occupations (by a factor of 1.3), while they would increase the most for workers in health diagnosing occupations (by a factor of 3.9). Counterfactual aggregate occupational mobility rates would be approximately 1.5 times higher if there were no costs to switching occupations.

The rest of the paper is organized as follows. Section 4.2 describes how the Eaton and Kortum (2002) gravity model may be adapted to think about flows of workers across occupations and the barriers to occupational mobility. Section 4.3 describes the empirical strategy and the proxies considered for the barriers to switching occupations. Section 4.4 describes the data. Section 4.5 presents the findings of the paper, while Section 4.6 concludes and suggests possibilities for future work.

4.2 Model

The model is an adaptation of the Eaton and Kortum (2002) model, modified to think about flows of workers across occupations and the barriers to occupational mobility. The framework involves a static, partial equilibrium model with perfect information.

4.2.1 Setup

There is a (finite) set of occupations given by $j \in \{1, 2, ..., N\}$, with a large number of employers in each occupation. Each occupation j produces a final good which has a price given by p_j. There is a continuum of homogeneous workers of measure 1, indexed by i. They only differ in terms of the occupation that they start the period in (which is exogenously pre-determined). A worker's occupation at the beginning of the period is indexed by k.

Workers maximize their current wages by choosing their optimal occupation. That is, at the beginning of the period, they consider the wages that they would receive in every possible occupation and select into the occupation where their potential wage is highest for that period.[66]

[66]Extending the model in order to have forward looking agents who maximize a present discounted value of current and expected future wages remains as future work. If the continuation value in each occupation were specific *either* to the occupation *or* to the individual, all of the analysis would still go through. However, within the setup of the model, if workers maximize present discounted values of wages, the continuation values

4.2.2 Wages and Productivity Draws

Workers' wages are equal to the value of their marginal product. Productivity is perfectly observable and labor is the only input in production.

Worker i's potential wage in occupation j is given by:

$$w_j(i) = p_j f\left[X(i)\right]\left(\frac{z_j(i)}{d_{kj}}\right) \tag{4.1}$$

p_j is the price of the output being produced in occupation j. $X(i)$ are a set of individual characteristics that augment productivity in all occupations. They may include variables such as education or overall work experience, which reflect general human capital.[67] $z_j(i)$ is an occupation-specific productivity draw for the individual. The process by which individuals draw productivity will be described in detail shortly. d_{kj} represents the cost of switching between the worker's initial occupation k and the potential occupation j. Assume $d_{kk} = 1$ (staying in the same occupation is costless) and $d_{kj} > 1$ for all $j \neq k$. The idea is that workers lose a fraction of their productivity if the occupation j that they choose in period t is different from the occupation k that they started the period in. The cost is assumed to last only for one period. Intuitively, if an individual is a newcomer to occupation j ($j \neq k$), then he must spend a fraction of his time training or learning to do his job. Therefore, in the period when he is a newcomer, his productivity per unit of time is $f\left[X(i)\right](z_j(i)/d_{kj}) < f\left[X(i)\right] z_j(i)$.

For each occupation j, individuals draw $z_j(i)$ from a Fréchet (or Type II extreme-value) distribution. One can think of individuals as receiving offers from several different potential employers in each occupation. The only offer the individuals will consider will be the highest offer in each occupation. Thus, the set of 'relevant' offers for each occupation are the collection of maxima across firms for each individual. The distribution of the maxima of a set of draws can converge to one of only three distributions, one of which is Fréchet (the distribution assumed here).[68]

The Fréchet distribution is given by:

$$z_j \sim F_j(z) = e^{-T_j z^{-\theta}} \tag{4.2}$$

Individuals sample all occupations at the beginning of the period (including the occupation they start in), by making a productivity draw for

would be specific to the occupation-individual combination.

[67] $X(i)$ does not include any characteristics that vary across occupations for the same individual, such as occupational tenure.

[68] See Eaton and Kortum (2002), footnote 14, and references therein.

each one. They then compare potential wages in all occupations. Based on their productivity draws, and the transition costs d_{kj}, they choose optimally whether to switch occupations, as well as which particular occupation to switch to.[69]

All individuals make draws from the same set of distributions $F_j(z)$, regardless of their current occupation or their individual characteristics. Therefore, for each occupation, all individuals have the same ex-ante expected productivity draw.[70]

Note that taking logs of Equation (4.1), we have:

$$\ln w_j(i) = \ln p_j + \ln f\left[X(i)\right] - \ln d_{kj} + \ln z_j(i) \tag{4.3}$$

This is similar to the wage equations commonly estimated in the literature, with $\ln p_j$ being an occupation wage premium and $\ln f\left[X(i)\right]$ being the return to a set of observable characteristics. This equation, however, also includes the switching cost $\ln d_{kj}$, and has an error term that is extreme-value distributed.

4.2.3 Flows of Workers Across Occupations

This section analyzes the implications of the model in terms of flows of workers across occupations. For individual i (who starts in occupation k), the probability that his wage in occupation j is above some level w is given by:

$$Pr\left[w_j(i) > w\right] = 1 - F_j\left(\frac{w d_{kj}}{p_j f\left[X(i)\right]}\right) \tag{4.4}$$

$$= 1 - e^{-T_j d_{kj}^{-\theta}(p_j f[X(i)])^{\theta} w^{-\theta}} \tag{4.5}$$

The probability that individual i draws a wage below w in every occu-

[69]There are no search frictions, and productivity draws are considered guaranteed job offers from an employer in an occupation, so the individual faces no uncertainty when choosing which occupation to switch to.

[70]The decision to switch occupations, however, will be influenced by the individual's current occupation through the 'penalty' for switching (d_{kj}). Individual characteristics $X(i)$ will not influence the switching decision because they augment wages in the same proportion in all potential occupations.

pation *other than* j is:

$$\begin{aligned} Pr\left[w_s(i) \leq w, \forall s \neq j\right] &= \prod_{s \neq j} F_s\left(\frac{w d_{ks}}{p_s f\left[X(i)\right]}\right) \\ &= \prod_{s \neq j} e^{-T_s d_{ks}^{-\theta}(p_s f[X(i)])^{\theta} w^{-\theta}} \end{aligned} \tag{4.6}$$

Individual i will optimally choose to switch to occupation j, given his current occupation k, if j offers him the highest potential wage among all possible occupations. The probability that this happens is denoted by $\pi_{kj}(i)$ and is given by:

$$\begin{aligned} \pi_{kj}(i) &\equiv Pr\left[w_j(i) \geq \max_s \{w_s(i)\}\right] \\ &= \int_0^{\infty} Pr\left[w_s(i) \leq w, \forall s \neq j\right] \cdot dPr\left[w_j(i) \leq w\right] \\ &= \frac{T_j d_{kj}^{-\theta} p_j^{\theta}}{\sum_{s=1}^{N} T_s d_{ks}^{-\theta} p_s^{\theta}} \end{aligned} \tag{4.7}$$

Intuitively, j will be the best choice for individual i if j offers him a wage w while all other occupations offer him a wage below w. Integrating this over all possible values of w gives us the probability that i switches from k to j, $\pi_{kj}(i)$. Note that this allows the possibility that $j = k$, i.e. the optimal choice may involve staying in the current occupation.

Taking the ratio of $\pi_{kj}(i)$ and $\pi_{kk}(i)$, based on Equation (4.7), and using the fact that $d_{kk} = 1$, we have:

$$\frac{\pi_{kj}(i)}{\pi_{kk}(i)} = \frac{T_j d_{kj}^{-\theta} p_j^{\theta}}{T_k p_k^{-\theta}} \tag{4.8}$$

Note that none of the terms on the right-hand-side are individual-specific, so the i index can be dropped, and, taking logs of the ratio, we have:

$$\ln \frac{\pi_{kj}}{\pi_{kk}} = \ln T_j + \theta \ln p_j - \ln T_k - \theta \ln p_k - \theta \ln d_{kj} \tag{4.9}$$

Because there is a continuum of individuals in each occupation, and because all individuals in occupation k have the same probability of switching to occupation j and all make independent draws from the productivity distribution, π_{kj} will be equal to the fraction of k-workers who switch to j, that is:

$$\pi_{kj} = \frac{sw_{kj}}{N_k} \tag{4.10}$$

where sw_{kj} is the total number of switchers from k to j and N_k is the size of occupation k (at the start of the period).

Therefore, Equation 4.9 can be rewritten in terms of worker flows, leading to a gravity-type equation:

$$\ln \frac{sw_{kj}}{sw_{kk}} = \ln T_j + \theta \ln p_j - \ln T_k - \theta \ln p_k - \theta \ln d_{kj} \tag{4.11}$$

This is the key equation of the model. It relates the flows of workers between occupations to a set of occupation-specific characteristics (T_j, T_k, p_j, p_k), and to the cost of switching (d_{kj}).[71]

4.3 Empirical Implementation

The goal is to estimate Equation (4.11) empirically, in order to understand the costs of switching between occupations (d_{kj}), and the factors that affect this variable. One of the main factors that might influence the cost of switching between two particular occupations, is the degree of similarity in the task content of the two occupations. To test this, I follow previous literature in characterizing occupations through a vector of skill or task characteristics and construct a measure of distance between occupation pairs based on this information. The distance measure reflects the degree of dissimilarity in the mix of tasks performed in the two occupations, and will therefore serve as an 'intensive margin' description of the occupational transition.[72] If human capital is task specific, one would expect that the cost of switching will increase with the task distance of the switch.[73]

The distance measure I consider is the log of an Euclidean distance. Let x_k^a be the importance level of dimension a (a certain ability) in occupation

[71]For the purposes of this paper, I do not close the model and do not consider the determinants of the prices p_j.

[72]Previous literature that has characterized occupations according to the tasks performed by the workers is numerous. Some examples are Autor et al. (2003), Ingram and Neumann (2006), Poletaev and Robinson (2008), and Peri and Sparber (2009). Gathmann and Schonberg (2010) were the first to suggest the concept of 'task distance'.

[73]This parallels the traditional use of gravity models in the trade literature, which consider the role of geographic distance between countries. Here, geographic distance is replaced with distance in terms of the heterogeneity in task content.

k, and x_j^a the analogous measure for occupation j. The Euclidean distance is measured as:

$$eucl\_dist_{kj} = \sqrt{\sum_{a=1}^{A} \left(x_k^a - x_j^a\right)^2} \tag{4.12}$$

where A is the total number of dimensions being considered. I work with a log-distance measure, which I define as:

$$dist_{kj} = \ln(1 + eucl\_dist_{kj}) \tag{4.13}$$

In addition to the task distance between occupations, I consider two other factors that might affect the cost of switching. The first one is whether k and j belong to the same major occupational group (1-digit occupation code). The second is whether k and j share the same type of 'main' task. Following Autor et al. (2003), main tasks can be either non-routine interactive, non-routine cognitive, routine cognitive, routine manual or non-routine manual, and are based on the dimension (out of these five) which has the highest importance score for each occupation. These additional factors might account for further similarities in the task content of occupations that are not captured fully by the distance measure.

Finally, I also allow for an error term ϵ_{kj} in d_{kj}, which captures barriers to occupational mobility arising from all other factors. This leads to the following specification for $\ln d_{kj}$:

$$\ln d_{kj} = \beta_1 dist_{kj} + \beta_2 common\_maj_{kj} + \beta_3 common\_tint_{kj} + \epsilon_{kj} \tag{4.14}$$

where $dist_{kj}$ is the distance measure, $common\_maj_{kj}$ is the indicator for common major occupational group, and $common\_tint_{kj}$ is the indicator for common main task. ϵ_{kj} is assumed to be a normally distributed random variable.

Using the specification for d_{kj}, and defining $D_j \equiv \ln T_j + \theta \ln p_j$, the gravity equation (4.11) becomes:

$$\ln \frac{sw_{kj}}{sw_{kk}} = D_j - D_k - \theta\beta_1 dist_{kj} - \theta\beta_2 common\_maj_{kj} - \theta\beta_3 common\_tint_{kj} - \theta\epsilon_{kj} \tag{4.15}$$

Empirically, the left-hand-side of the equation can be measured using data on worker flows across occupations, while D_j and D_k can be captured through occupation fixed effects. To keep the estimation flexible, I allow source and destination occupation effects to vary. That is, the main estimating equation is given by:

$$\ln \frac{sw_{kj}}{sw_{kk}} = D_j - S_k - \theta\beta_1 dist_{kj} - \theta\beta_2 common\_maj_{kj} - \theta\beta_3 common\_tint_{kj} - \theta\epsilon_{kj} \tag{4.16}$$

where D is a destination occupation fixed effect, S is a source occupation fixed effect, and observations are at the occupation pair level.

4.4 Data

I measure flows of workers across occupations using data from the Current Population Survey (CPS). The CPS is the source of many official Government statistics in the United States. It is administered by the Bureau of the Census under the auspices of the Bureau of Labor Statistics (BLS). In recent years, a nationally representative sample of about 65,000 households has been interviewed monthly.[74] The March Supplement of the CPS directly asks individuals about their current occupation, and the occupation they worked in during the previous year.[75] I use annual data from the March CPS from 1983 to 2002. Over this period, the CPS used a consistent occupational coding system.[76] The sample is restricted to adults aged 18 to 65 who are not in the military.

I use occupation codes at the 2-digit Census code level.[77] More detailed occupation systems (i.e. 3-digit codes) provide a level of aggregation that is too fine to observe significant flows of workers across particular occupation pairs. Meanwhile, a higher level of aggregation (1-digit level) creates groups that are too coarsely grouped together.

[74] See Moscarini and Thomsson (2007) for some of the advantages of the CPS. See also Kambourov and Manovskii (2004) for some of its limitations. The main disadvantage of the CPS for the purposes of this paper is that it is address-based, thus generating attrition when geographical mobility occurs, and geographical mobility is potentially correlated with occupational mobility.

[75] I use the commercially available version of the data from Unicon Research Corporation. Many of the CPS variables can also be obtained from IPUMS (King et al., 2010).

[76] Over time, the CPS has used several different occupational coding schemes (Census codes or 'COC'). 1968-70 was coded in 1960 COC; 1971-82 in 1970 COC; 1983-91 in 1980 COC; 1992-2002 in 1990 COC; 2003-09 in 2000 COC. The change between 1980 and 1990 COC, however, was minor and does not create an important break in the data. I take care of this minor change by converting the occupation codes to the standardized codes from Meyer and Osborne (2005). Peter Meyer's code is available at `http://econterms.net/pbmeyer/research/occs/remapjob.do`.

[77] See Appendix Table C.1 for the definition and content of each 2-digit code.

To build annual flows of workers across occupations, I use interviewee's responses to the questions of their occupation last year (previous occupation), and occupation last week (current occupation). The flow of switchers between occupation k and occupation j (sw_{kj}) is defined as the number of respondents that report occupation k as their previous occupation, and occupation j as their current occupation.

The skill characterization of occupations uses data from the Revised 4th Edition of the Dictionary of Occupational Titles (ICPSR, 1991). The DOT provides precise measures of the different abilities that are required in different occupations, as well as the different work activities performed by job incumbents. The dimensions along which the DOT dataset characterizes occupations include complexity of work, General Education Development (GED), specific vocational preparation requirements, aptitudes, temperaments and physical demands, among others (ICPSR, 1981). The choice of the relevant dimensions with which to characterize occupations in order to construct a distance measure is somewhat arbitrary. I choose the three GED variables and the eleven aptitudes from the 1991 DOT as the relevant dimensions. Table 4.1 provides examples of the DOT task vectors for four particular occupations. Distance measures are constructed based on these dimensions as described in the previous section. Appendix C.1 explains how the CPS data on worker flows and the DOT data on occupational characteristics and distance are merged with each other.

4.5 Results

4.5.1 Gravity Equation Estimation

This section presents the results from estimating Equation (4.16) using the CPS data on worker flows and the proxies for mobility costs described above. Note that the left-hand-side of Equation (4.16) is the log of the ratio of switchers to stayers for each occupation pair. In practice, there are several occupation pairs for which the flow of workers is equal to zero in the data. I deal with these occupation pairs with zero-flows in three alternative ways. In Column (1) of Table 4.2, I present results from the estimation of Equation (4.16) when the ratios that are equal to zero are changed to a small positive number (0.00001). In Column (2), I instead drop all observations with zero flows from the estimation. Finally, in Column (3) I change the zeros in the same way as for Column (1), and estimate a Tobit regression, where the data is assumed to be censored at the log of 0.00001. All specifications include source and destination occupation fixed effects, which are not shown in the

Table 4.1: Examples of DOT-91 scores

Occupation	Math & Comp Scientists	Retail Salesp.	Private household clean & serv	Construct trades
Reasoning	1.51	-0.20	-1.19	-0.38
Math	2.00	0.09	-1.04	-0.51
Language	1.63	0.01	-0.95	-0.70
Intelligence	1.76	-0.32	-0.77	-0.46
Verbal	1.81	-0.19	-0.70	-0.61
Numerical	3.55	-0.01	-0.73	-0.46
Spatial	0.41	-0.43	-0.69	0.02
Form Percep	0.05	-0.54	-0.82	0.07
Clerical Per	2.36	0.02	-0.77	-0.70
Motor Coord	-0.98	-0.64	-1.07	0.49
Finger Dext	-0.63	-0.30	-0.71	0.11
Manual Dext	-1.04	-0.82	-0.24	0.69
Eye-Hand-Foot	-0.39	-0.20	0.08	1.33
Color Discrim	-0.77	0.03	-0.63	0.01

Table, but will be discussed shortly. The regressions also include a full set of year dummies.

The Table shows that the effect of task distance on worker flows is negative and significant, suggesting that task distance is an important component of the cost of occupational mobility. Meanwhile, being in the same major occupational category or sharing the same main task significantly increase worker flows, as they reduce the cost of transitioning. The estimated coefficients presented in the Table are of $-\theta\beta$ (see Equation (4.16)), so the interpretation of the magnitudes is left to the next subsection, where I disentangle these two components.

The estimated coefficients on the source and destination occupation dummies for the OLS specification excluding zero flows (Column (2) of Table 4.2) are plotted against each other in Figure 4.1. Each dot corresponds to a particular 2-digit occupation (labeled in the graph with the corresponding occupation code, the definitions of which are in the Appendix). The excluded category in the regression, 'Retail and other salespersons', is given a value of zero for both its source and destination effects.

The model predicts that if an occupation is an attractive destination (for example, because of a high value of T_j relative to other occupations), it

Table 4.2: Estimated coefficients on 'gravity-type' equation, 1983-2002

	OLS All	OLS Non-zero	Tobit
	(1)	(2)	(3)
logd91	-.863	-.400	-1.366
	(.018)***	(.008)***	(.032)***
common-maj	1.024	.505	1.606
	(.038)***	(.015)***	(.064)***
common-tint	.224	.033	.435
	(.029)***	(.012)***	(.049)***
Obs.	41624	22682	41624
R^2	.455	.591	

Note: The Table presents the results of estimating Equation (4.16). The dependent variable is $\ln(sw_{kj}/sw_{kk})$. All specifications include source and destination occupation dummies, as well as year dummies.

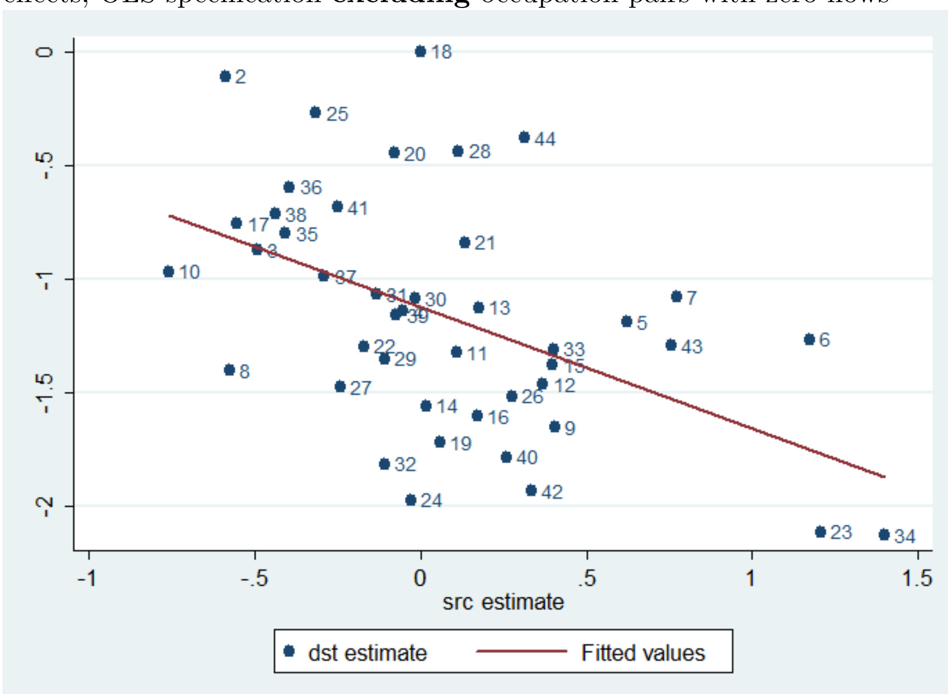
will have high entry rates (a high destination fixed effect) and low exit rates (a low source fixed effect). In fact, as Equation (4.15) shows, it predicts a perfect negative correlation between source and destination fixed effects. Figure 4.1 shows that the estimated source and destination occupation fixed effects from the specification that excludes occupation pairs with zero flows are negatively correlated, as the theory predicts. The Figure plots the line of best fit. Its slope is significantly different from zero (point estimate equal to -0.48, standard error 0.11), but also significantly different from -1.

For specifications that include occupation pairs with zero flows, however, estimated source and destination occupation fixed effects become strongly positively correlated. Figure 4.2 plots the estimated occupation fixed effects from the OLS specification that includes pairs with zero flows (Column (1) of Table 4.2). The strong positive correlation means that occupations that have large exit rates also tend to attract a large share of switchers from other occupations, suggesting a high degree of churning. Figure 4.3 plots the same estimated source occupation fixed effects against occupation size, and shows the strong correlation between these two variables. Large occupations tend to have high rates of both entry and exit.

4.5.2 Estimation of θ

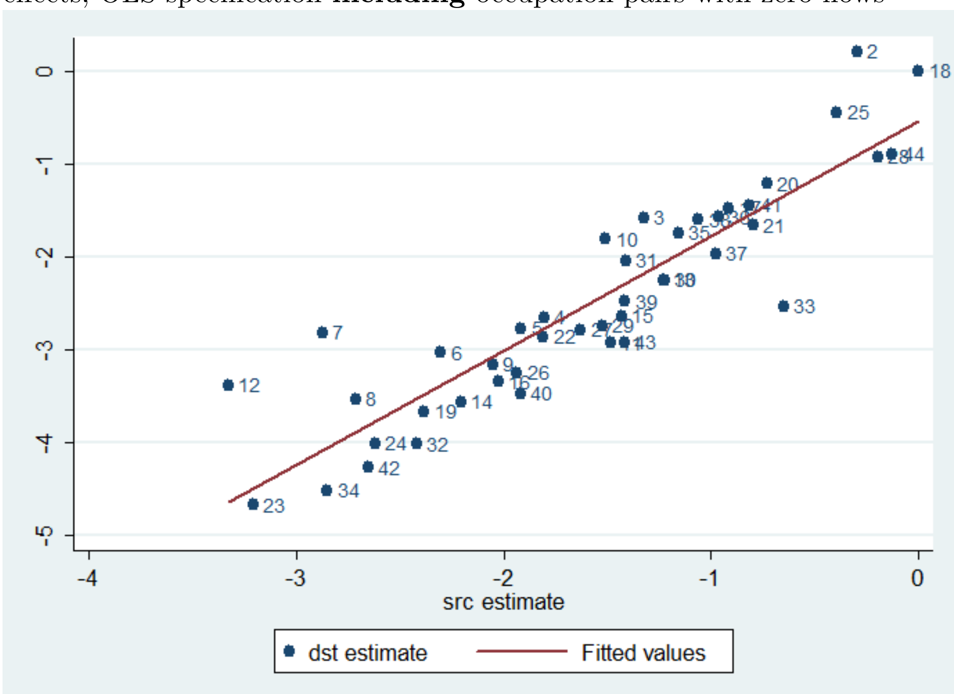
Equation (4.16) shows that the estimated coefficients presented in Table 4.2 are of $-\theta\beta$. An estimate of θ is needed in order to get β, the marginal effects

Figure 4.1: Scatter plot of estimated source and destination occupation fixed effects, OLS specification **excluding** occupation pairs with zero flows



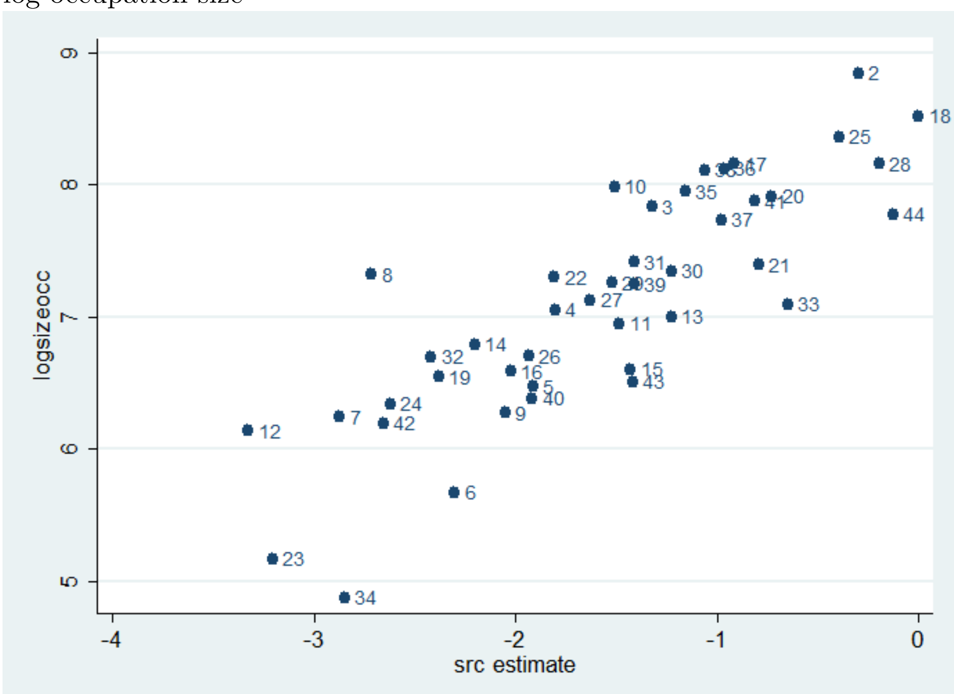
Note: The numbers indicate the corresponding 2-digit occupation. See Appendix for the definition of each occupation code. 2-digit occupation code 1 is excluded because it is very small (in terms of number of workers).

Figure 4.2: Scatter plot of estimated source and destination occupation fixed effects, OLS specification **including** occupation pairs with zero flows



Note: The numbers indicate the corresponding 2-digit occupation. See Appendix for the definition of each occupation code. 2-digit occupation code 1 is excluded because it is very small (in terms of number of workers).

Figure 4.3: Scatter plot of estimated source occupation fixed effects from the OLS specification **including** occupation pairs with zero flows against log occupation size



Note: The numbers indicate the corresponding 2-digit occupation. See Appendix for the definition of each occupation code. 2-digit occupation code 1 is excluded because it is very small (in terms of number of workers).

of the variables of interest on d_{kj}, the cost of switching occupations.

In order to estimate θ, I take advantage of the properties of the extreme value distribution that is assumed in the model, which lead to a simple relationship between the distribution of wages predicted by the model and the parameter of interest θ.

First, the extreme value distribution in Equation (4.2) has mean $T_j^{1/\theta}\Gamma\left(1-1/\theta\right)$, where Γ is the Gamma function. Its log has a Gumbel distribution, with standard deviation equal to $\pi/(\theta\sqrt{6})$.[78]

Now, note that the (ex-post) wage for an individual starting in occupation k is also drawn from an extreme value distribution. The probability that an individual ends up with a wage below or equal to w is equal to the probability that his potential wage in *all* possible occupations is below or equal to w. That is:

$$\begin{aligned} Pr(w(i) \leq w|k) &= Pr\left[w_j(i) \leq w \ \forall j|k\right] \\ &= \prod_{j=1}^{N} F_j\left(\frac{wd_{kj}}{p_j f\left[X(i)\right]}\right) \\ &= e^{-\left(\sum_{j=1}^{N} T_j d_{kj}^{-\theta}(p_j f[X(i)])^{\theta}\right)w^{-\theta}} \end{aligned} \tag{4.17}$$

This distribution has mean $\left[\sum_{j=1}^{N} T_j d_{kj}^{-\theta}(p_j f\left[X(i)\right])^{\theta}\right]^{(1/\theta)}\Gamma\left(1-1/\theta\right)$, and its log has standard deviation $\pi/(\theta\sqrt{6})$.[79]

Therefore, the standard deviation of ex-post log wages for a set of individuals starting in occupation k with common demographic characteristics (common $X(i)$), is a function of θ. The standard deviation for the entire set of individuals starting in k (unconditional on demographics) will be at least as large as the (common) standard deviation within each sub-group. In other words, the standard deviation of ex-post log wages for the entire set of individuals starting in occupation k will be bounded below by $\pi/(\theta\sqrt{6})$. That is, $\sigma_k \geq \pi/(\theta\sqrt{6})$, so $\theta \geq \pi/(\sigma_k\sqrt{6})$.

[78] The Gumbel distribution has a CDF given by: $F(x) = exp(-e^{-(x-\mu)/\beta})$, with standard deviation $(\pi\beta)/\sqrt{6}$. In the case of the log of the productivity draws from the extreme value distribution in Equation (4.2), we have that: $Pr(\ln z_j(i) \leq z) = Pr(z_j(i) \leq e^z) = F_j(e^z) = exp(-T_j e^{-\theta z})$, which is a Gumbel distribution with standard deviation $\pi/(\theta\sqrt{6})$.

[79] Note from the previous footnote that the standard deviation of the log of productivity does not depend on T_j. In the case of the distribution of log wages, T_j would be replaced by $\left(\sum_{j=1}^{N} T_j d_{kj}^{-\theta}(p_j f\left[X(i)\right])^{\theta}\right)$. The standard deviation remains independent of this multiplicative constant.

Table 4.3: Summary statistics of estimates of $\underline{\theta}$ and θ for occupations with more than 100/20 reports over the period 1983-2002

Sample:	Full	Restricted
Estimate of:	$\underline{\theta}$	θ
10th Percentile	1.5536	1.7874
Median	1.9866	2.4461
Mean	2.0275	2.5145
90th Percentile	2.4946	3.3255

Note: The table is based on the distribution of the estimated values of θ as described in the text. The full sample includes estimated values of θ for each occupation with at least 100 reports. The restricted sample includes estimated values of θ for each education group by initial occupation, restricted to workers aged 25-30, and to groups with at least 20 earnings reports. Estimates are annual over the period 1983-2002.

It is possible therefore to get a lower-bound estimate of θ by calculating $\underline{\theta} = \pi/(\sigma_k\sqrt{6})$, where σ_k is the standard deviation of log wages for the set of individuals starting the period in occupation k. I use this property to get an estimate of $\underline{\theta}$ from the data. For each previous occupation in each year I will have an estimate per year of $\underline{\theta}$.

In order to get a more precise estimate of θ, I also estimate the standard deviation within groups with common demographics. In particular, I split the sample into age-education ('skill') bins and focus on the groups of workers aged 25 to 30. As they are at the start of their working life, it is reasonable to expect low heterogeneity in terms of the characteristics that determine their wages. I therefore get estimates of θ based on the standard deviation of wages for people aged 25 to 30, by initial occupation and separately for workers with high school or less, and workers with some college and college.

Figure C.1 in the Appendix shows histograms of the estimated values of θ, for the entire sample, using occupations with more than 100 valid reports, and for the restricted sample, using occupation-skill groups with at least 20 observations, for the period 1983-2002. Table 4.3 shows the corresponding summary statistics.

4.5.3 Implied Mobility Costs

In Table 4.4, I use the median estimate of θ from the restricted sample to calculate the estimated effects of the different variables on the iceberg cost.

Table 4.4: Estimated effects on occupational transition costs for each of the specifications in Table 4.2

Variable	Implied β			Percentage effect on d_{kj}		
	OLS All	OLS Non-zero	Tobit	OLS All	OLS Non-zero	Tobit
logd91	0.353	0.163	0.558	42.187	16.307	74.555
common_maj	-0.419	-0.207	-0.657	-34.212	-18.662	-48.135
common_tint	-0.091	-0.013	-0.178	-8.737	-1.340	-16.291

Note: Percentage effect on d_{kj} refers to the effect of a 1 standard deviation increase in distance in the corresponding sample (all switches or non-zero only), and of a change from 0 to 1 for the other two variables.

The results in this Table are based on the estimated coefficients in Table 4.2. The Table shows that the impact of task distance on the cost of switching is substantial. The implied coefficients from the OLS specification excluding zero flows suggest that a one standard deviation increase in distance increases the cost of switching occupations by 16.3% if the occupations are in different major occupational groups and do not share the same type of main task. The effect is reduced to 15% if the occupation pair shares the same type of main task. If the occupations are in the same major occupation group, however, the implied cost of the transition is negligible under this specification.

Table 4.5 shows the estimated transition barriers for some commonly observed occupational switches. The estimates use the implied β from the OLS specification which excludes transitions with zero flows, and take into account the values taken on by the distance measure, and the indicators for common major occupational group and common main task for each occupation pair. Consider workers switching between food service occupations and retail occupations. The estimation results imply a cost (wage penalty) of 33% ($\hat{d}_{kj} = 1.33$). For a small number of occupation pairs, $\hat{d}_{kj}$ is below 1, which would mean that the cost of switching is negative. In the table, this is the case for workers switching between 'other administrative support occupations' and 'information and records processing occupations'. Only a small fraction of occupation pairs are estimated to have negative transition costs (less than 3.4% of occupation pairs have $\hat{d}_{kj} < 1$, and less than 1.2% have $\hat{d}_{kj} < 0.9$).

Table 4.5: Estimated barriers to occupational transition for some commonly observed transitions (OLS specification excluding occupation pairs with zero flows)

Occupations		$\hat{d}_{kj}$
Other executive, administrators, and managers	Retail and other salespersons	1.40
Food service occupations	Retail and other salespersons	1.33
Other admin support occ, including clerical	Retail and other salespersons	1.17
Health assessment and treating occupations	Health technologists and technicians	1.02
Sales supervisors and sales reps, finance and business	Retail and other salespersons	1.02
Other admin support occ, including clerical	Info and records processing, except financial	0.84

4.5.4 Counterfactual Experiments

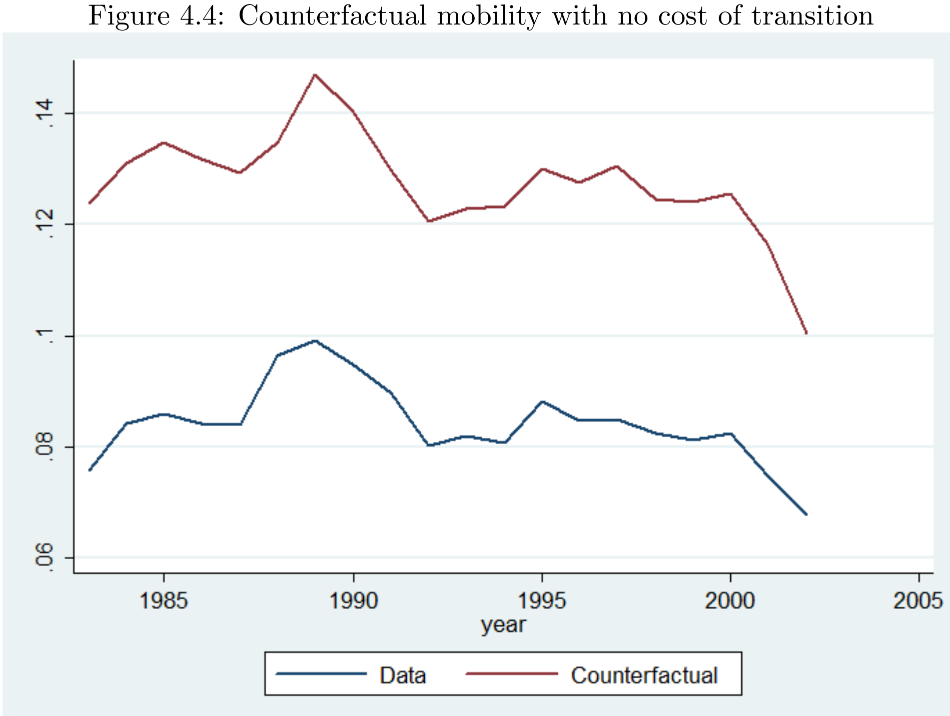
The final experiment performed in the paper consists of estimating counterfactual mobility rates if transition costs were zero. I calculate the fraction of workers that would be predicted to switch out of occupation k, if switching were costless. For each occupation, I define $ratio_k$ as $\widehat{Mob}_k^{CF}/\widehat{Mob}_k$, where $\widehat{Mob}_k$ is total fitted mobility out of occupation k, given the estimated cost of switching, and $\widehat{Mob}_k^{CF}$ is total counterfactual mobility with no transition costs. $\widehat{Mob}_k^{CF}$ is driven entirely by the estimated source and destination occupation fixed effects. Table 4.6 presents estimates of $ratio_k$ for the occupations with the lowest and the highest ratios (again for the OLS specification excluding zero flows). The ratios range from 1.3 to close to 4. The results of the counterfactual experiment suggest that the fraction of workers switching out of management occupations would increase by 30% if managers could switch costlessly. For health diagnosing occupations, the costs of switching are estimated to be much higher, so the fraction of switchers would be almost four times higher if they could switch at no cost.

Figure 4.4 shows aggregate counterfactual mobility rates with no cost of transition. This figure confirms the substantial role of mobility costs. If they did not exist, aggregate occupational mobility rates would be approximately 1.5 times as high as they are in the data.

Table 4.6: Ratio of counterfactual mobility to fitted mobility for selected occupations

Rank	Occ Code	Occ Name	$ratio_k$
1	3	Management related occupations	1.31
2	23	Office machine and communication equip operators	1.54
3	43	Helpers, construction and production occ	1.56
4	24	Mail and message distributing	1.58
41	6	Natural scientists	2.88
42	20	Secretaries, stenographers, and typists	3.06
43	4	Engineers, architects	3.22
44	7	Health diagnosing occupations	3.92

Figure 4.4: Counterfactual mobility with no cost of transition



4.6 Conclusion

This paper provides a new approach to the analysis of the costs of occupational mobility. A version of the gravity model commonly used in the trade literature is adapted in order to analyze flows of workers across occupation pairs, and the barriers that they face. In the model, workers make productivity draws from a set of extreme value distributions for each potential occupation, and choose optimally which occupation to work in, based on their productivity draw and the costs of transition from their current occupation.

The empirical implementation focuses on the role of task distance (degree of dissimilarity in the mix of task requirements) on the cost of switching occupations. This variable is found to play an important role in explaining the barriers to occupational mobility, with a one standard deviation of task distance increasing the cost of switching occupations by at least 16%, all else equal. Meanwhile, being in the same major occupational group plays a significant role in reducing mobility costs, in ways not captured by the distance measure. Finally, counterfactual occupational mobility rates would be substantially higher if workers were able to switch occupations at no cost.

The paper may be extended in several ways. First, the model abstracts from the role of occupational tenure in the accumulation of (occupation-specific) human capital. Appendix C.2 describes how tenure can be added to the theoretical framework. Second, the specification of the mobility costs imposes perfect symmetry between the costs of switching from occupation k to occupation j (d_{kj}), and the corresponding switch in the opposite direction (d_{jk}). In practice, occupational transitions involve different types of skill upgrading (for example due to career progression), or skill downgrading (for example due to unexpected job displacement). The costs of occupational 'upgrades' might be very different from those of the corresponding occupational 'downgrades', and this is something not currently captured in the specification of transition costs. Third, an analysis of the extent to which coding error biases the results is important. Presumably a fraction of the worker flows occurring between similar (low distance) occupations could be attributable to coding error. It would be important to find ways to address this issue. Finally, a more substantial extension of the model would entail considering a dynamic setting where workers maximize a present discounted value of lifetime utility, rather than current wage only.

Chapter 5

Conclusion

This dissertation provides an extensive analysis of the importance of the task content of occupations. Chapter 2 provides a theoretical and an empirical framework to study the impacts of task-replacing technologies on individual workers. Chapters 3 and 4 develop a measure of task distance and analyze its relevance in explaining the wage changes associated with occupational transitions of individual workers, as well as the observed aggregate flows of workers across occupations.

The findings in Chapter 2 bridge the literatures on individual-level occupational mobility and on aggregate-level polarization and technological change. The results presented in that Chapter help interpret many of the findings from the occupational mobility literature within the broader context of technical change and labor market polarization. Routine-biased technical change is theoretically consistent as the driving force behind increasing occupational mobility, selection on ability for occupational switchers, and changes in occupation wage premia over time.

The dissertation also contributes to the literature on the human capital costs associated with occupational mobility. By adding a concept of distance, Chapters 3 and 4 go beyond the traditional analysis that considers all occupational transitions as being equal. The findings suggest that there is substantial heterogeneity in the costs associated with different types of occupational switches. Switches between occupations with very different task content may be much more costly in terms of human capital losses than switches between more similar occupations, but only for certain subsets of workers. The evidence presented in the dissertation suggests that switching costs act as an important barrier to occupational mobility, but only across and not within broad occupation groups. These findings are informative in order to understand the costs faced by workers who experience negative shocks in their current occupations, whether it is due to business cycle shocks or longer term trends such as technological change. The results also shed light, however, on the limitations of using task distance as a generalized measure for the cost of switching occupations, as this may not be appropriate in certain cases.

There are many potential avenues for future work based on the findings in this dissertation. The findings in Chapters 3 and 4 may serve as a stepping stone for further analyses on the human capital implications of occupational switches. The specification of the distance measure considered in this dissertation (Euclidean distance) is perfectly symmetric, implying that the effect of switching from occupation k to occupation j (d_{kj}), and the corresponding switch in the opposite direction (d_{jk}) are identical. In practice, occupational transitions involve different types of skill upgrading (for example due to career progression), or skill downgrading (for example due to unexpected job displacement). The costs of occupational 'upgrades' might be very different from those of the corresponding occupational 'downgrades', and this is something not currently captured in the specification of transition costs. Relatedly, the analysis presented here does not make a distinction between voluntary and involuntary occupational switches. Another direction for future research would be to separately explore the role played by different specific task components in order to understand along which dimensions it is particularly costly for workers to transition. These decompositions would help us understand in which particular cases it is appropriate to think of task distance as a measure of transition costs, or how best to define this distance measure.

It remains as future work to use the findings in the dissertation in order to build an integrated dynamic model with partially-transferable human capital, where workers maximize a present discounted value of lifetime utility by making occupational choices. This would make it possible to think about the causes and implications of the differences in mean distance for different cross-sections of the population, and would generate more precise measures of the aggregate human capital costs of occupational mobility. Such a model would also be useful for an analysis of an economy's reaction to different types of shocks to particular tasks. For example, in a model with specific human capital, if there is a technology shock such as the one considered in Chapter 2, a worker who switches occupations would be trading off an initial loss of specific human capital against the access to an occupation that is performing relatively better. This type of consideration could explain the findings in Chapter 2 regarding the wage changes for switchers to non-routine manual occupations: They experience an initial wage cut but subsequently recover from it, as wages in non-routine manual occupations are growing faster than they do in routine ones. Finally, it also remains as future work to formally quantify the extent to which the polarization phenomenon discussed in Chapter 2 has occurred through job displacement or through changing employment patterns of new entrants to the labor force.

Bibliography

Acemoglu, D. and D. Autor (2011), "Skills, tasks and technologies: Implications for employment and earnings." *Handbook of Labor Economics*, 4, 1043–1171.

Andrews, M., T. Schank, and R. Upward (2006), "Practical fixed-effects estimation methods for the three-way error-components model." *Stata Journal*, 6, 461.

Autor, D. and D. Dorn (2009), "This job is "getting old": Measuring changes in job opportunities using occupational age structure." *American Economic Review*, 99, 45–51.

Autor, D.H., L.F. Katz, and M.S. Kearney (2006), "The Polarization of the US Labor Market." *The American Economic Review*, 96, 189–194.

Autor, D.H., L.F. Katz, and M.S. Kearney (2008), "Trends in US wage inequality: revising the revisionists." *The Review of Economics and Statistics*, 90, 300–323.

Autor, D.H., F. Levy, and R.J. Murnane (2003), "The Skill Content of Recent Technological Change: An empirical exploration." *Quarterly Journal of Economics*, 118, 1279–1333.

Bartel, A., C. Ichniowski, and K. Shaw (2007), "How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills." *The Quarterly Journal of Economics*, 122, 1721–1758.

Berman, E., J. Bound, and Z. Griliches (1994), "Changes in the Demand for Skilled Labor within US Manufacturing: Evidence from the Annual Survey of Manufactures." *The Quarterly Journal of Economics*, 109, 367.

Berman, E., J. Bound, and S. Machin (1998), "Implications of skill-biased technological change: International evidence." *Quarterly Journal of Economics*, 113, 1245–1279.

Boudarbat, B., W.C. Riddell, and T. Lemieux (2010), "The Evolution of the Returns to Human Capital in Canada, 1980–2005." *Canadian Public Policy*, 36, 63–89.

Costinot, A. and J. Vogel (2010), "Matching and inequality in the world economy." *Journal of Political Economy*, 118, 747–786.

Cragg, M.I. and M. Epelbaum (1996), "Why has wage dispersion grown in Mexico? Is it the incidence of reforms or the growing demand for skills?" *Journal of Development Economics*, 51, 99–116.

Dustmann, C., J. Ludsteck, and U. Schönberg (2009), "Revisiting the German Wage Structure." *Quarterly Journal of Economics*, 124, 843–881.

Dustmann, C. and C. Meghir (2005), "Wages, experience and seniority." *The Review of Economic Studies*, 72, 77–108.

Eaton, J. and S. Kortum (2002), "Technology, geography, and trade." *Econometrica*, 70, 1741–1779.

Firpo, S., N. Fortin, and T. Lemieux (2011), "Occupational tasks and changes in the wage structure." *University of British Columbia*.

Gathmann, C. and U. Schonberg (2010), "How General is Human Capital? A Task-Based Approach." *Journal of Labor Economics*, 28, 1–49.

Gibbons, R., L. Katz, T. Lemieux, and D. Parent (2005), "Comparative advantage, learning, and sectoral wage determination." *Journal of Labor Economics*, 23, 681–724.

Goldin, C.D. and L.F. Katz (2008), "Long-run changes in the wage structure: Narrowing, widening, polarizing." *Brookings Papers on Economic Activity*, 2007, 135–165.

Goos, M. and A. Manning (2007), "Lousy and lovely jobs: The rising polarization of work in Britain." *The Review of Economics and Statistics*, 89, 118–133.

Goos, M., A. Manning, and A. Salomons (2009), "Job polarization in Europe." *American Economic Review*, 99, 58–63.

Gregory, R. and F.G. Vella (1995), "Real Wages, Employment, and Wage Dispersion in US and Australian Labor Markets." *NBER Chapters*, 205–226.

Groes, F., P. Kircher, and I. Manovskii (2009), "The U-Shapes of Occupational Mobility." *University of Pennsylvania mimeo*.

Grossman, G.M. and E. Rossi-Hansberg (2008), "Trading tasks: A simple theory of offshoring." *The American Economic Review*, 98, 1978–1997.

Gürtzgen, N. (2009), "Rent-sharing and collective bargaining coverage: Evidence from linked employer–employee data." *The Scandinavian Journal of Economics*, 111, 323–349.

Harris, D.N. and T.R. Sass (2010), "Teacher training, teacher quality and student achievement." *Journal of Public Economics*.

Heathcote, J., F. Perri, and G.L. Violante (2010), "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967-2006." *Review of Economic Dynamics*, 13, 15–51.

Heckman, James J. and Yona Rubinstein (2001), "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *The American Economic Review*, 91, 145–149.

Hummels, David, Rasmus Jorgensen, Jakob R. Munch, and Chong Xiang (2011), "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data." Working Paper 17496, National Bureau of Economic Research.

ICPSR, 1981 (1981), *Dictionary of Occupational Titles (DOT): Part I - Current Population Survey, April 1971, Augmented with DOT Characteristics, and Part II - Fourth Edition Dictionary of DOT Scores for 1970 Census Categories [Computer File]. Study Nr. 07845*. National Academy of Sciences - Committee on Occupational Classification and Analysis, Washington DC: U.S. Department of Commerce, Bureau of the Census [Producer], Ann Arbor, MI: Inter-university Consortium for Political and Social Research [Distributor].

ICPSR, 1991 (1991), *Dictionary of Occupational Titles (DOT): Revised Fourth Edition, 1991 [Computer file]. Study Nr. 06100. United States Department of Labor, United States Employment Service, and the North Carolina Occupational Analysis Field Center*. United States Department of Labor, United States Employment Service, and the North Carolina Occupational Analysis Field Center, Washington, DC: United States Department of Labor, United States Employment Service, and Raleigh, NC: North Carolina Occupational Analysis Field Center [Producer], 1991. Ann

Arbor, MI: Inter-university Consortium for Political and Social Research [Distributor], 1994.

Ingram, B.F. and G.R. Neumann (2006), "The returns to skill." *Labour Economics*, 13, 35–59.

Juhn, C., K.M. Murphy, and B. Pierce (1993), "Wage inequality and the rise in returns to skill." *Journal of Political Economy*, 410–442.

Jung, J. and J. Mercenier (2010), "Routinization-biased technical change, globalization and labor market polarization: Does theory fit the facts?" *Working Papers ERMES, U. Pantheon-Assas (Paris 2)*.

Kambourov, G. and I. Manovskii (2004), "A Cautionary Note on Using (March) CPS Data to Study Worker Mobility." *manuscript, University of Pennsylvania, Philadelphia*.

Kambourov, G. and I. Manovskii (2008), "Rising occupational and industry mobility in the United States: 1968-1997." *International Economic Review*, 49, 41–79.

Kambourov, G. and I. Manovskii (2009a), "Occupational mobility and wage inequality." *Review of Economic Studies*, 76, 731–759.

Kambourov, G. and I. Manovskii (2009b), "Occupational specificity of human capital." *International Economic Review*, 50, 63–115.

Katz, L.F. and D.H. Autor (1999), "Changes in the wage structure and earnings inequality." *Handbooks of Labor Economics*, 3, 1463–1558.

Katz, L.F. and K.M. Murphy (1992), "Changes in relative wages, 1963–1987: supply and demand factors." *The Quarterly Journal of Economics*, 107, 35.

King, Miriam, Steven Ruggles, Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick (2010), "Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]." Technical report, University of Minnesota, Minneapolis, MN.

Krueger, A.B. (1993), "How computers have changed the wage structure: Evidence from microdata, 1984-1989." *The Quarterly Journal of Economics*, 33–60.

Krusell, P., L.E. Ohanian, J.V. Ríos-Rull, and G.L. Violante (2000), "Capital-skill complementarity and inequality: A macroeconomic analysis." *Econometrica*, 68, 1029–1053.

Lazear, Edward P. (2003), "Firm-specific human capital: A skill-weights approach." Working Paper 9679, National Bureau of Economic Research.

Lefter, A. and B.M. Sand (2011), "Job Polarization in the US: A Reassessment of the Evidence from the 1980s and 1990s." *University of St. Gallen, School of Economics and Political Sciences, Department of Economics, Economics Working Paper Series.*

Lemieux, T. (2008), "The changing nature of wage inequality." *Journal of Population Economics*, 21, 21–48.

Martins, P.S. and L.D. Opromolla (2009), "Exports, imports and wages: Evidence from matched firm-worker-product panels." *IZA Discussion Papers.*

Meyer, PB and AM Osborne (2005), "Proposed Category System for 1960-2000 Census Occupations." *Bureau of Labor Statistics Working Paper*, 383.

Milkman, R. and R.E. Dwyer (2002), "Growing Apart: The "New Economy" and Job Polarization in California, 1992-2000." *The State of California Labor 2002.*

Moscarini, G. and K. Thomsson (2007), "Occupational and Job Mobility in the US." *Scandinavian Journal of Economics*, 109, 807–836.

Moscarini, G. and F.G. Vella (2008), "Occupational Mobility and the Business Cycle." *NBER Working Papers.*

Murphy, K.M. and F. Welch (1993), "Occupational change and the demand for skill, 1940-1990." *The American Economic Review*, 83, 122–126.

Neal, D. (1995), "Industry-specific human capital: Evidence from displaced workers." *Journal of labor Economics*, 13, 653–677.

Nordhaus, W.D. (2007), "Two centuries of productivity growth in computing." *Journal of Economic History*, 67, 128.

Parent, D. (2000), "Industry-specific capital and the wage profile: Evidence from the national longitudinal survey of youth and the panel study of income dynamics." *Journal of Labor Economics*, 18, 306–323.

Peri, G. and C. Sparber (2009), "Task specialization, immigration, and wages." *American Economic Journal: Applied Economics*, 1, 135–169.

Poletaev, M. and C. Robinson (2008), "Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984-2000." *Journal of Labor Economics*, 26, 387–420.

Robinson, C. (2010), "Occupational Mobility, Occupation Distance and Basic Skills: Evidence from Job-Based Skill Measures."

Robinson, C. (2011), "Occupational mobility, occupation distance and specific human capital." *University of Western Ontario, CIBC Centre for Human Capital and Productivity Working Papers*.

Schank, Thorsten, Claus Schnabel, and Joachim Wagner (2007), "Do exporters really pay higher wages? First evidence from German linked employeremployee data." *Journal of International Economics*, 72, 52 – 74.

Sullivan, P. (2010), "Empirical evidence on occupation and industry specific human capital." *Labour Economics*, 17, 567–580.

Topel, R. (1991), "Specific capital, mobility, and wages: Wages rise with job seniority." *Journal of Political Economy*, 99, 145.

Wooldridge, J.M. (2002), *Econometric analysis of cross section and panel data*. The MIT press.

Yamaguchi, Shintaro (2012), "Tasks and heterogeneous human capital." *Journal of Labor Economics*, 30, pp. 1–53.

Appendix A

Appendices to Chapter 2

A.1 Proposition Proofs

A.1.1 Proof of Proposition 1

First, consider the signs of the partial derivatives of Equations (2.14) to (2.16). Using the fundamental theorem of calculus, we have that:

$$man'(z_0) \equiv \frac{\partial man(z_0)}{\partial z_0} = \varphi_{man}(z_0)g(z_0) > 0 \tag{A.1}$$

$$rt_0(z_0, z_1) \equiv \frac{\partial rt(z_0, z_1)}{\partial z_0} = -\varphi_{rt}(z_0)g(z_0) < 0 \tag{A.2}$$

$$rt_1(z_0, z_1) \equiv \frac{\partial rt(z_0, z_1)}{\partial z_1} = \varphi_{rt}(z_1)g(z_1) > 0 \tag{A.3}$$

$$cog'(z_1) \equiv \frac{\partial cog(z_1)}{\partial z_1} = -\varphi_{cog}(z_1)g(z_1) < 0 \tag{A.4}$$

Intuitively, increases in z_0 (given z_1) increase the measure of workers performing non-routine manual tasks, and decrease the measure of workers performing routine tasks. Increases in z_1 (given z_0) increase the measure of workers performing routine tasks, and decrease the measure of workers performing non-routine cognitive tasks.

Meanwhile, based on the assumptions on absolute and comparative advantage across skill groups (Equation (2.5)), we have that α_0 and $\tilde{\alpha}_1$ are decreasing functions in their respective arguments. That is:

$$\alpha_0'(z_0) = \frac{d \ln \varphi_{man}}{dz_0} - \frac{d \ln \varphi_{rt}}{dz_0} < 0 \tag{A.5}$$

$$\tilde{\alpha}_1'(z_1) = \frac{\varphi_{rt}(z_1)}{\varphi_{cog}(z_1) + \varphi_{rt}(z_1)} \left[\frac{d \ln \varphi_{rt}}{dz_1} - \frac{d \ln \varphi_{cog}}{dz_1} \right] < 0 \tag{A.6}$$

It follows that $\Delta > 0$ and the signs of the general equilibrium effects of a change in $\ln \kappa_{rt}$ are as indicated in Proposition 1.

A.1.2 Proof of Proposition 2

C_{man} is constant, as it is equal to p_1 which is normalized to 1. So d $d \ln C_{man}/d \ln \kappa_{rt} = 0$.

From Equation (2.7),

$$\begin{aligned} \ln C_{rt} &= \ln\left(\frac{\varphi_{man}(z_0)}{\varphi_{rt}(z_0)}\right) + \ln C_{man} \\ &= \alpha_0(z_0) + \ln C_{man} \end{aligned}$$

So:

$$\begin{aligned} \frac{d \ln C_{rt}}{d \ln \kappa_{rt}} &= \alpha_0'(z_0)\frac{dz_0}{d \ln \kappa_{rt}} + \frac{d \ln C_{man}}{d \ln \kappa_{rt}} \\ &= \alpha_0'(z_0)\frac{dz_0}{d \ln \kappa_{rt}} \end{aligned}$$

From Proposition 1 and its proof, $dz_0/d \ln \kappa_{rt} > 0$ and $\alpha_0'(z_0) < 0$. Therefore, $d \ln C_{rt}/d \ln \kappa_{rt} < 0$.

From Equation (2.8),

$$\begin{aligned} \ln C_{cog} &= \ln\left(\frac{\varphi_{rt}(z_1)}{\varphi_{cog}(z_1)}\right) + \ln C_{rt} \\ &= \alpha_1(z_1) + \ln C_{rt} \end{aligned}$$

So:

$$\frac{d \ln C_{cog}}{d \ln \kappa_{rt}} = \alpha_1'(z_1)\frac{dz_1}{d \ln \kappa_{rt}} + \frac{d \ln C_{rt}}{d \ln \kappa_{rt}}$$

Given the assumptions on comparative advantage, $\alpha_1'(z_1) < 0$. From Proposition 1, $dz_1/d \ln \kappa_{rt} < 0$, so the first term is positive. However, $d \ln C_{man}/d \ln \kappa_{rt} < 0$, so the sign of $d \ln C_{cog}/d \ln \kappa_{rt}$ is ambiguous.

From Equation (2.8),

$$\ln\left(\frac{C_{cog}}{C_{rt}}\right) = \alpha_1(z_1)$$

So:

$$\frac{d \ln(C_{cog}/C_{rt})}{d \ln \kappa_{rt}} = \alpha_1'(z_1)\frac{dz_1}{d \ln \kappa_{rt}}$$

which is unambiguously positive and implies:

$$\frac{d \ln C_{cog}}{d \ln \kappa_{rt}} > \frac{d \ln C_{rt}}{d \ln \kappa_{rt}}$$

A.1.3 Proof of Proposition 3

The wage change for a worker of skill level z is given by:

$$\frac{d\ln w(z)}{d\ln \kappa_{rt}} = \frac{d\ln C_{j'}}{d\ln \kappa_{rt}} + \mathbb{1}(j \neq j'|z)\left[\ln C_{j'} - \ln C_j + \ln \varphi_{j'}(z) - \ln \varphi_j(z)\right] \tag{A.7}$$

where j denotes the occupation that a worker of skill level z optimally chooses before the change in κ_{rt}, and j' indicates his optimal choice after the shock. $\mathbb{1}(j \neq j'|z)$ is an indicator function equal to 1 if $j \neq j'$ for a worker of skill level z.

For workers who do not switch occupations, $j = j'$, and the wage change induced by the change in κ_{rt} is equal to the change in the wage per efficiency unit C_j in their occupation.

For workers who do switch occupations, the wage change is equal to the change in $C_{j'}$ (change in wage per efficiency unit in their new occupation) plus the difference between the wage they would have received before the shock in occupation j' and the wage they were receiving in occupation j.

Based on the results from Proposition 1, and defining the new cutoff skill levels (after the shock to κ_{rt}) as z_0' and z_1', we have that:

$$\mathbb{1}(j \neq j'|z) = \begin{cases} 0 & \text{if } z_{min} \leq z < z_0 \text{ or } z_0' \leq z < z_1' \text{ or } z_1 \leq z \leq z_{max} \\ 1 & \text{if } z_0 \leq z < z_0' \text{ or } z_1' \leq z < z_1 \end{cases}$$

For non-switchers, the proof of Proposition 3 follows directly from Proposition 2. For switchers, the proof is as follows.

Consider first switchers to non-routine manual, that is, workers with ability $z \in [z_0, z_0')$. The first term, $d\ln C_{man}/d\ln \kappa_{rt}$ is zero. Now, recall that before the shock to κ_{rt}, all workers with $z \in [z_0, z_1)$ were optimally sorting into routine jobs, meaning that, given their skill level and the pre-shock equilibrium values of C_{man} and C_{rt}, the wage they would have earned in non-routine manual jobs was lower than the wage they were earning in routine jobs. Formally, this means: $\ln C_{rt} + \ln \varphi_{rt}(z) > \ln C_{man} + \ln \varphi_{man}(z)$ for all workers with ability $z \in [z_0, z_0')$. It follows that the wage change for workers in this ability range is unambiguously negative.

Now consider switchers to non-routine cognitive occupations, that is, workers with ability $z \in [z_1', z_1)$. First consider a worker of ability z_1'. As z_1' is the new cutoff, it must be the case that this worker is indifferent between working in a routine or a non-routine cognitive job, or in other words, the wage change he experiences if he switches or he stays should be the same.

This implies:

$$\begin{aligned}\frac{d\ln w(z_1')}{d\ln\kappa_{rt}} &= \frac{d\ln C_{cog}}{d\ln\kappa_{rt}} + \ln C_{cog} + \ln\varphi_{cog}(z_1') - \big(\ln C_{rt} + \ln\varphi_{rt}(z_1')\big)\\ &= \frac{d\ln C_{rt}}{d\ln\kappa_{rt}}\\ &< 0\end{aligned}$$

where the second line is the wage change if the worker had not switched occupations, which must be equal to the first line due to the indifference condition between switching and staying.

This shows that a worker of ability z_1' will experience a wage cut equal to that experienced by stayers in routine occupations. Now consider a worker of ability z_1. Before the shock these workers would have been indifferent between routine and non-routine cognitive jobs, meaning that $\ln C_{cog} + \ln\varphi_{cog}(z_1) = \ln C_{rt} + \ln\varphi_{rt}(z_1)$. It follows then that:

$$\begin{aligned}\frac{d\ln w(z_1)}{d\ln\kappa_{rt}} &= \frac{d\ln C_{cog}}{d\kappa_{rt}} + \ln C_{cog} + \ln\varphi_{cog}(z_1) - (\ln C_{rt} + \ln\varphi_{rt}(z_1))\\ &= \frac{d\ln C_{cog}}{d\ln\kappa_{rt}}\\ &> \frac{d\ln C_{rt}}{d\ln\kappa_{rt}}\end{aligned}$$

The wage change for workers of ability z_1 is unambiguously greater than the wage change for stayers in routine jobs. The conclusion is that, among the workers who switch to non-routine cognitive, the wage change is increasing in the worker's ability level, and their wage change is greater or equal to that experienced by workers staying in routine occupations.

A.2 Model Extension: Two-Dimensional Skills

This section describes how the model can be extended to account for two-dimensional skills that are rewarded differently in different occupations. It is shown that, by imposing certain restrictions on the variances of those skills in the population and the differences in the returns across occupations, the predictions of the model are still valid in expectations.

In particular, assume that workers are endowed with a certain level of cognitive skills z_i^{cog} and manual skills z_i^{man}. The marginal productivity of these skills varies across occupations. For simplicity, assume that only cognitive skills are productive in non-routine cognitive occupations, and only manual skills are productive in non-routine manual occupations. Both types of skills are productive in routine occupations.

The marginal productivity of an individual with skills $\{z_i^{cog}, z_i^{man}\}$ is given by:

$$\ln \varphi_j(z_i) = \begin{cases} b_{man}^{man} z_i^{man} & \text{in non-routine manual jobs} \\ b_{rt}^{man} z_i^{man} + b_{rt}^{cog} z_i^{cog} & \text{in routine jobs} \\ b_{cog}^{cog} z_i^{cog} & \text{in non-routine cognitive jobs} \end{cases}$$

where the subscript on b indicates the occupation and the superscript indicates the type of skill.

The predictions of the model in terms of the sorting patterns and the switches induced by routinization will still be true if the following conditions hold:

$$Cov\left\{w_{cog}(z_i^{man}, z_i^{cog}) - w_{rt}(z_i^{man}, z_i^{cog}), w_{rt}(z_i^{man}, z_i^{cog})\right\} > 0 \tag{A.8}$$

$$Cov\left\{w_{man}(z_i^{man}, z_i^{cog}) - w_{rt}(z_i^{man}, z_i^{cog}), w_{rt}(z_i^{man}, z_i^{cog})\right\} < 0 \tag{A.9}$$

where $w_j(z_i^{man}, z_i^{cog})$ are the wages received in occupation j by a worker of skills $\{z_i^{cog}, z_i^{man}\}$. The covariances imply that routine workers with higher wages will in expectation be the ones with more to gain from switching to non-routine cognitive, while the workers with relatively lower wages will in expectation be the ones with more to gain from switching to non-routine manual.

Assume that the endowments of the two types of skills are independently distributed in the population, so that $Cov(z_i^{man}, z_i^{cog}) = 0$. The variances of each of the two types of skills are denoted σ_{man}^2 and σ_{cog}^2. Using these assumptions, along with the equation for log wages (2.6) and the assumption for $\ln \varphi_j(z_i)$, Equation (A.8) implies:

$$Cov\left\{\theta_{cog} - \theta_{rt} + (b_{cog}^{cog} - b_{rt}^{cog}) z_i^{cog} - b_{rt}^{man} z_i^{man}, \theta_{rt} + b_{rt}^{cog} z_i^{cog} + b_{rt}^{man} z_i^{man}\right\} > 0$$
$$(b_{cog}^{cog} - b_{rt}^{cog}) b_{rt}^{cog} \sigma_{cog}^2 - (b_{rt}^{man})^2 \sigma_{man}^2 > 0$$

That is:

$$(b_{cog}^{cog} - b_{rt}^{cog}) b_{rt}^{cog} \sigma_{cog}^2 > (b_{rt}^{man})^2 \sigma_{man}^2 \tag{A.10}$$

And Equation (A.9) implies:

$$Cov\left\{\theta_{man} - \theta_{rt} + (b_{man}^{man} - b_{rt}^{man}) z_i^{man} - b_{rt}^{cog} z_i^{cog}, \theta_{rt} + b_{rt}^{cog} z_i^{cog} + b_{rt}^{man} z_i^{man}\right\} < 0$$
$$(b_{man}^{man} - b_{rt}^{man}) b_{rt}^{man} \sigma_{man}^2 - (b_{rt}^{cog})^2 \sigma_{cog}^2 < 0$$

That is:

$$(b^{man}_{man} - b^{man}_{rt}) b^{man}_{rt} \sigma^2_{man} < (b^{cog}_{rt})^2 \sigma^2_{cog} \tag{A.11}$$

Equations (A.10) and (A.11) provide restrictions on the dispersion of skills in the population, and on the returns to skills in the different occupations. As long as these two equations hold, the predicted patterns of occupational switching described in the paper go through.

Note that the restrictions allow for returns to manual skills to be largest in non-routine manual occupations (that is $(b^{man}_{man} - b^{man}_{rt}) > 0$), as long as the variance of cognitive skills in the population is large enough relative to the variance of manual skills.

A.3 Grouping of Occupation Codes

Table A.1 describes the mapping of 3-digit occupation codes into the broad categories used in the main specification in the paper. The mapping is based on aggregating 3-digit codes into 1-digit categories, and then labeling them according to their main task (Acemoglu and Autor, 2011).

For the alternative classification of occupations used in Section 2.7.3, I directly use task data from the Dictionary of Occupational Titles (DOT), and follow a procedure similar to Autor and Dorn (2009) to assign occupations to the broad task categories (non-routine manual, routine and non-routine cognitive). I use data from the 4th Edition of the DOT, which was published in 1977 and is available in electronic format through the Interuniversity Consortium for Political and Social Research (ICPSR, 1981). The DOT provides precise measures of the different abilities that are required in different occupations, as well as the different work activities performed by job incumbents. The DOT-77 has its own coding scheme, which is much more disaggregated than the Census Occupation Codes (COC) used in the PSID. To aggregate to the 1970-COC level, I follow Autor et al. (2003). I use the April 1971 CPS Monthly File, in which experts assigned individuals both with 1970-COC and DOT-77 codes. Using the CPS sampling weights, I calculate means of each DOT task measure at the 1970-COC occupation level. Each DOT score is rescaled to have a (potential) range from zero to 10. I then generate an index of relative routine task intensity for each occupation j (RTI_j) as follows:

$$RTI_j = \frac{rt_j}{max\{nr\_cog_j, nr\_man_j\}} \tag{A.12}$$

Table A.1: Occupation code groupings

Task label	Occupations included	3-digit Census Codes 1970-COC	2000-COC
nr_cog	Professional, technical and kindred workers	001-195	
	Professional and related occupations		100-354
	Managers, officials and proprietors, except farm	201-245	
	Management, business and financial occupations		001-095
	Managers of retail and non-retail sales workers		470-471
routine	Sales workers, except managers	260-285	472-496
	Clerical and kindred workers	301-395	
	Office and administrative support occupations		500-593
	Craftsmen, foremen and kindred workers	401-575	
	Operatives, except transport	601-695	
	Laborers, except farm	740-785	
	Construction and extraction occupations		620-694
	Installation, maintenance and repair occupations		700-762
	Production occupations		770-896
	Transport equipment operatives	701-715	
	Transportation and material moving occupations		900-975
nr_man	Service workers	901-984	360-465
Not classified	Members of armed forces	600	984
	Farmers, farm managers, farm laborers, farm foremen	801-824	
	Farming, fishing and forestry occupations		600-613

Note: The 1970 Census Occupation Codes (COC) were used in the PSID up to 2001. Since 2003, the 2000 coding system has been used. Task labels are based on Acemoglu and Autor (2011). Occupation code groupings and details on the 3-digit codes can be found in the Working Paper version of Kambourov and Manovskii (2008) and on the IPUMS website (King et al., 2010): See `http://usa.ipums.org/usa/volii/97occup.shtml` for the 1970 codes and `http://usa.ipums.org/usa/volii/00occup.shtml` for the 2000 codes.

where, following Autor et al. (2003), rt_j is the mean score for 'Dealing with set limits, tolerances and standards' and 'Finger dexterity'; nr_cog_j is the mean score for 'Mathematics' and 'Direction, control and planning'; and nr_man_j is the score for 'Eye-Hand-Foot Coordination'. I attach the task measures to the 1970 Census (downloaded from David Autor's website) and label the occupations in the top third of the employment-weighted distribution of RTI as intensive in routine tasks.[80] Among the remaining occupations, I generate an index of relative non-routine cognitive task intensity (CTI_j) as follows:

$$CTI_j = \frac{nr\_cog_j}{nr\_man_j} \tag{A.13}$$

I label the occupations above the median of the employment-weighted distribution (among the remaining occupations) of CTI as intensive in non-routine cognitive tasks, and the remaining occupations as intensive in non-routine manual tasks. Once I have each 1970-COC labeled with its main task, I can attach these task labels to the PSID data up to 2001 (which is the period for which PSID occupation were coded using 1970-COC).

[80]The weights are equal to the product of the Census sampling weight, weeks worked last year and usual weekly hours.

Appendix B

Appendices to Chapter 3

B.1 Robustness Check - Evolution of Distance with O*Net Measures

In this Appendix, I describe the O*Net dataset and present a robustness check on the result regarding the increase in distance over time. I follow ideas developed by Kambourov and Manovskii (2008) in terms of dealing with coding error in the PSID.

B.1.1 O*Net Data

O*Net is the successor to the DOT. The first complete version of O*Net (O*Net 4.0) was published in 2002. In the DOT and in O*Net 4.0, occupational characteristics were based exclusively on analyst ratings. Since then, the O*Net Program began to update its data by conducting interviews of job incumbents. The latest version of O*Net was released in 2009 (O*Net 14.0). Its main source of data are the responses of job incumbents, although some data is still provided by occupational analysts. The collection of occupations included in O*Net has been periodically revised to include new occupations, and to incorporate changes in the nature of existing occupations. The data used in this paper is from O*Net 13.0 (2008).

O*Net Dimensions

The dimensions along which the O*Net Database characterizes occupations are the following:

- 52 Abilities, defined as attributes of an individual that influence work performance, e.g. “Oral Comprehension”, “Inductive Reasoning”, “Manual Dexterity”, “Stamina”. They are subdivided into:
  - Cognitive abilities
  - Psychomotor abilities
  - Physical abilities

  - Sensory abilities

- 16 Work Styles, defined as personal characteristics that can affect how well someone performs a job, e.g. "Persistence", "Leadership" or "Dependability".

- 41 Work Activities, defined as general types of job behaviors, e.g. "Analyzing Data or Information", "Making Decisions and Solving Problems", "Performing General Physical Activities", "Communicating with Supervisors, Peers, or Subordinates". Work activities are subdivided into:

  - Information input
  - Mental processes
  - Work output (physical, manual and technical activities)
  - Interacting with others

- Other dimensions: Basic skills, Cross-functional skills, Knowledge, Experience and training, Work context.

For each dimension, the question asked to job incumbents (or occupational experts) is: "How important is X to the performance of your job?", where X is a specific ability/work style/etc.[81] For each occupation, then, there is an importance score along each dimension. This score can range between 0 and 7.

B.1.2 Relevant Occupational Characteristics

The choice of the relevant dimensions with which to characterize occupations is somewhat arbitrary. I generate distance measures using different subsets of O*Net data. The subsets I consider are as described below:

1. All available information (distance measure: *d_all*): Occupations are characterized by a 109-dimensional vector that includes all abilities, work values and work activities.

2. All abilities (*d_aa*): Occupations are characterized by a 52-dimensional vector of ability scores.

[81] In some cases, an additional question is posed which asks "What level of X is needed to perform your job?". However, answers to the "importance" and to the "level" questions are very highly correlated (0.97 in the case of Abilities and 0.92 in the case of Work Activities in O*Net-2009), so the "level" data does not actually add any extra information.

3. All work activities (*d_at*): Occupations are characterized by a 41-dimensional vector of importance scores for tasks performed on the job.[82]

4. Selected abilities (*d_sa*): The previous measures are very rich, as they use a lot of information. However, it is implicitly weighting different subcategories in different ways. For instance, in the ability vector, 21 abilities are classified as cognitive abilities, while only 9 are classified as physical abilities, so the measure *d_aa* implicitly places a larger weight on cognitive rather than physical abilities. In order to consider measures of distance that include a more "balanced" mix of different types of abilities, I also develop a distance measure based on selected abilities. In particular, I choose two cognitive abilities (Deductive Reasoning, Mathematical Reasoning), two physical abilities (Manual Dexterity, Strength), and two non-cognitive abilities (Persistence, Social Orientation).[83]

5. Selected tasks (*d_st*): I similarly develop a measure of distance based on selected work activities. In particular, I choose two work activities related to mental processes (Analyzing data and information, Making decisions and solving problems), two physical and manual work activities (Performing general physical activities, Handling and moving objects), and two work activities related to interacting with others (Communicating with supervisors, peers or subordinates, Performing for or working directly with the public).

6. Most correlated abilities (*d_mca*): The previous measures (*d_sa* and *d_st*) incorporate arbitrary choices of abilities and tasks. In order to discipline the choice of abilities, I also develop a measure in which, for each of the four ability subcategories in O*Net (cognitive, psychomotor, physical, sensory) and for non-cognitive abilities (work styles), I choose the dimension that, for all occupations in the O*Net data, is most correlated with the other dimensions within its subcategory. In a sense this chooses the "most representative" ability within each subcategory. This procedure results in choosing the following abilities: Deductive reasoning (Cognitive), Control precision (Psychomotor), Stamina (Physical), Sound localization (Sensory), and Initiative (Non-cognitive).

[82]This would be the measure closest to Gathmann and Schonberg (2010)
[83]See Heckman and Rubinstein (2001) on the importance of non-cognitive skills.

7. Most correlated tasks (*d_mct*): I use the same procedure as the one for *d_mca* but using the different sub-categories of work activities. The resulting "representative" work activities are: Monitor processes, materials, or surroundings (Information input); Developing objectives and strategies (Mental processes); Repairing and maintaining mechanical equipment (Work output); Developing and building teams (Interacting with others).

B.1.3 Matching O*Net Data with the PSID

The classification of occupations in O*Net is based on SOC. The SOC classification can be converted to the 2000 Census classification using codes from the National Crosswalk Service Center.[84] To match the O*Net data at the 2000-COC level to the PSID data, I rely on a crosswalk developed by Meyer and Osborne (2005),[85] which proposes a series of standard occupation codes that can be linked to the different COCs.

Several SOC occupations receive the same 2000 COC, and several 2000 COC occupations in turn receive the same "standardized" (Meyer-Osborne) code. In these cases, I average the O*Net scores of all the SOC occupations that end up with the same standardized code. Since the standardized codes combine some occupations, this means that some occupational transitions observed in the PSID data with the 1970 Census codes are lost when they are converted to the standardized codes in order to merge them with O*Net information.

B.1.4 Adjustment for Coding Error

In this section I follow the methodology developed by Kambourov and Manovskii (2008) in adjusting for the prevalence of coding error in the PSID in the post-1981 period. In their paper they are interested in adjusting for the shift in the *fraction of switchers* from 1981 onwards. I apply the same method to adjust for the shift in measured *distance* in that period.

The idea is to rely on the higher level of accuracy of the occupational coding in the 1968-1980 period. The key assumption is that the incidence of the post-1981 coding error varies across age and education groups, but not over time. This assumption allows an estimation of the time trend of the mean distance of occupational switches by controlling for the change in the

[84] See `http://www.xwalkcenter.org/`.

[85] Peter Meyer's code is available at `http://econterms.net/pbmeyer/research/occs/remapjob.do`

coding procedure and its incidence on age-education groups. In particular, I run an OLS regression of the following form:

$$Dist = \beta_0 + \beta_1 Age + \beta_2 Age^2 + \beta_3 Time + \beta_4 Break + \beta_5 Educ + interactions \tag{B.1}$$

where all variables are indexed by i, t but the indices are omitted for notational simplicity. The dependent variable in Equation B.1, *Dist*, is the distance of the occupational transition. The control variables are age (in bins)[86], age squared, a time trend, *Break* (a dummy equal to 1 for 1981 onwards), interactions of time trend and *Break* with age and age squared, and interactions of all the above with an education dummy equal to one for people with some college or college education.

B.1.5 Findings

Figure B.1 shows the adjusted mean annual distances for occupational switchers using distance measure *d_all*. To generate adjusted distances, the implied effects of having $Break = 1$ from 1981 onwards are obtained from the estimation of Equation B.1. These implied effects are then subtracted from the observed and from the fitted values of distance in the sample.

The upward trend that is seen in the Figure is confirmed by computing the derivative of Equation B.1 with respect to the time trend. The value of this derivative, computed around the mean of all variables is positive and highly significant (p-value: [0.000]). To check the robustness of the finding of an increase in the mean distance of occupational switches, I estimate Equation B.1 for each of the alternative distance measures, i.e. All abilities, All work activities, Selected abilities and Selected tasks. The results for the adjusted values of mean distance using these alternative measures are plotted in Figure B.2. The figure shows that the mean distance of switches (conditional on switching) is increasing for all the distance measures being considered. The marginal estimated effect of the time trend is highly significant for all these measures.

I also estimate Equation B.1 for each of the sub-categories that are present in the O*Net database. In particular, as mentioned earlier, O*Net classifies its 52 abilities into 4 subgroups: cognitive, psychomotor, physical and sensory. Its 41 work activities are classified into 4 subgroups as well: information input, mental processes, work output, and interactive tasks.

[86] I create bins for the following age groups: 19-22, 23-25, 26-30, 31-35, 36-40, 41-45, 46-50, 51-60 and 61-70.

Figure B.1: Mean annual distance for occupational switchers adjusted for the change in the coding procedure

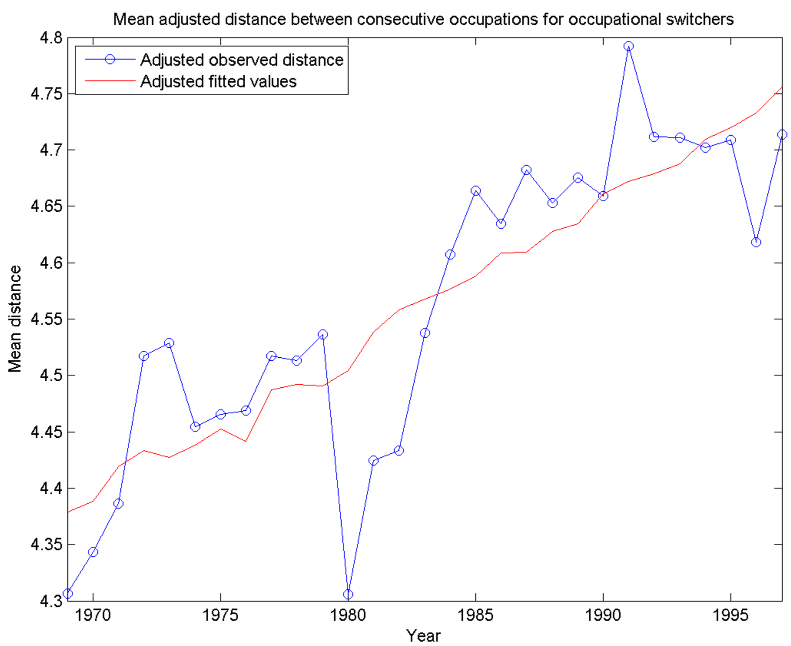


Figure B.2: Mean annual distance adjusted for coding error for alternative measures of distance

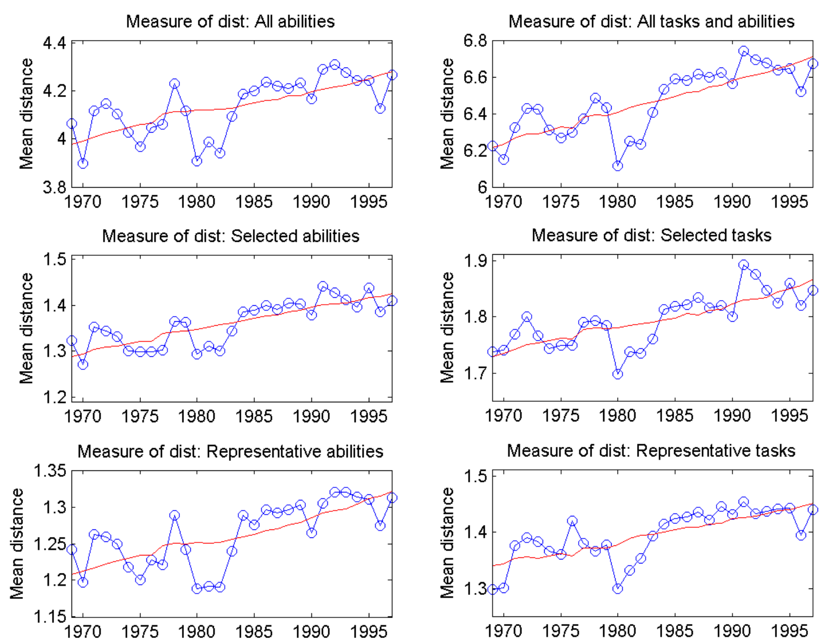
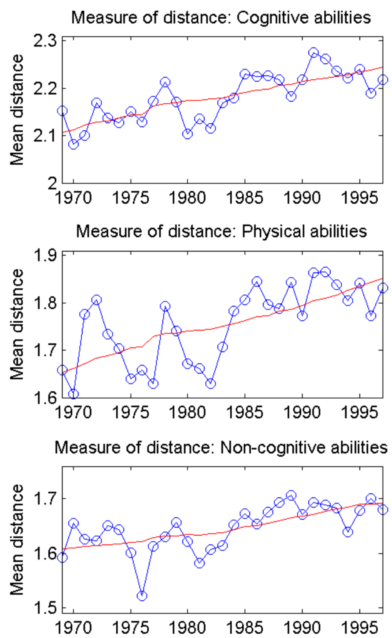


Figure B.3: Mean annual distance adjusted for coding error - Distance measures based on importance scores of different groups of abilities as indicated in each graph



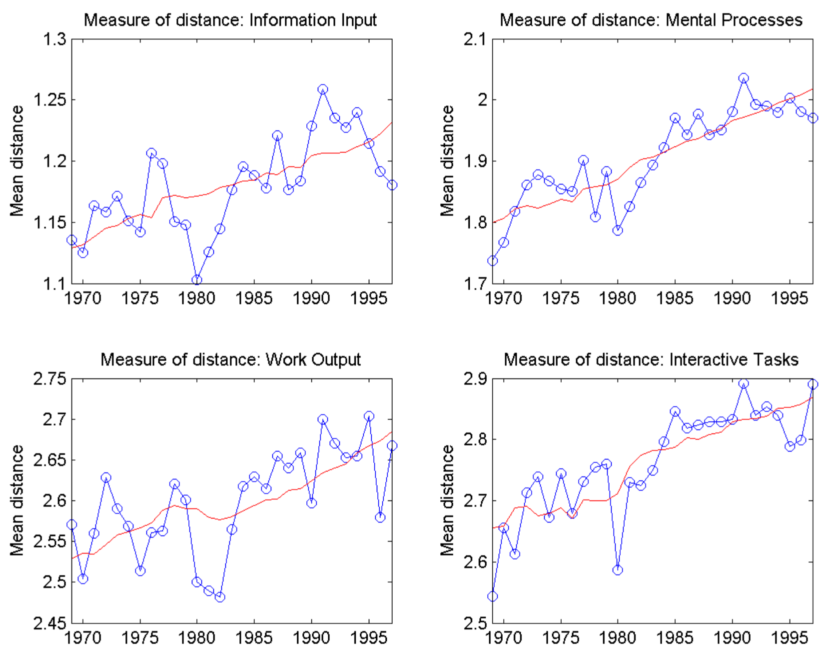
Derivatives with respect to time significant at 5% level for Sensory abilities; significant at 1% level for all other ability groups.

The graphs for the adjusted mean annual distance for the different ability categories, as well as the group of non-cognitive abilities (work styles) are presented in Figure B.3. The corresponding graphs for each of the task categories are in Figure B.4. The finding is that the increase in distance is a widespread phenomenon. The derivatives with respect to time for all of these categories are positive and significant at the 5% level or above.

In short, the increase in the mean distance of occupational switches (conditional on switching) is significant, robust to different distance specifications, and widespread among different sub-categories of abilities and work activities.[87]

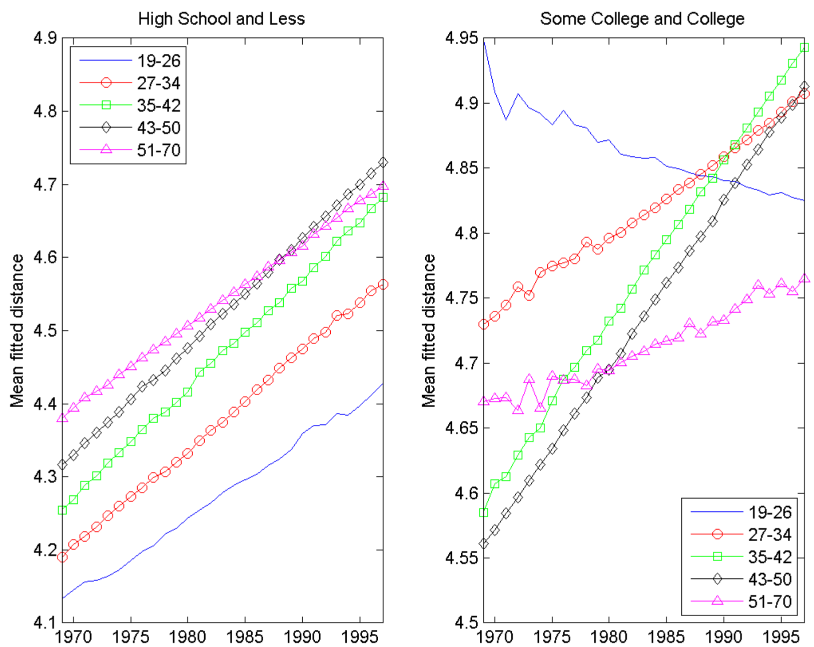
[87] I also check whether there is an increase in the dispersion of the observed distances. I do this by taking the distribution of the adjusted observed distance and calculating the ratio for each year of the 90th percentile to the median, and the ratio of the median to

Figure B.4: Mean annual distance adjusted for coding error - Distance measures based on importance scores of different groups of work activities as indicated in each graph



Derivatives with respect to time significant at 1% level for Mental processes and Work output; significant at 5% level for Information input and Interactive tasks.

Figure B.5: Adjusted mean distance of occupational switches (conditional on switching) by age and education groups



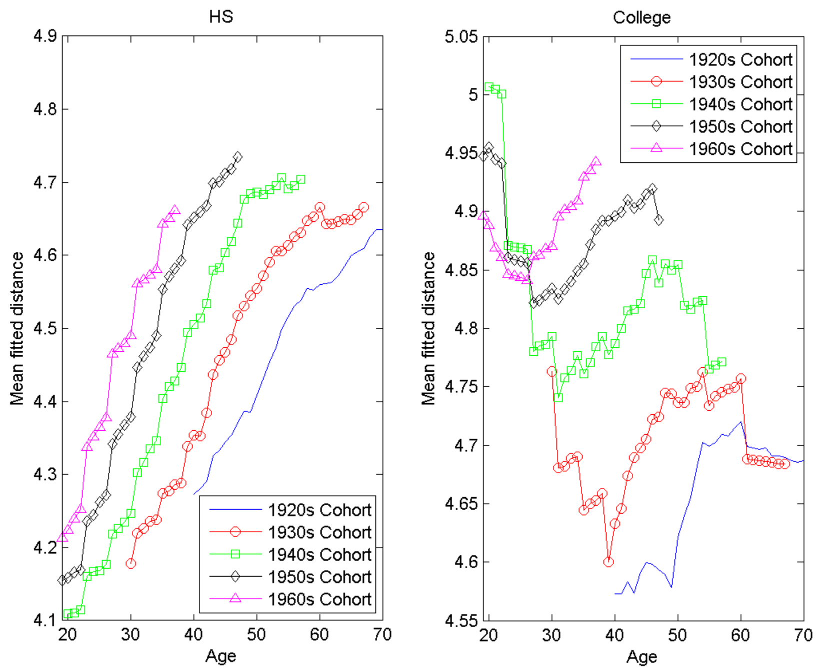
B.1.6 Patterns for Different Age and Education Groups

Next, I analyze the level and trend for the mean distance by age-education groups. By education, I divide the sample between workers with a high school degree or less, and workers with at least some college education. By age, I divide the sample into 5 groups: ages 19-26, 27-34, 35-42, 43-50 and 51-70. The results are presented in Figure B.5. I find that there is an important increase in the mean distance for all age groups among the less educated workers. For the more educated workers the increase over time is particularly strong for workers aged 35-50.

In order to address the changes across age-education groups more clearly, I divide the sample into cohorts, corresponding to different birth decades, from the 1920s to the 1960s. The results are in Figure B.6. Workers with

the 10th percentile. I do not find any patterns over time in these ratios, and therefore conclude that there are no patterns over time in the dispersion of observed distances.

Figure B.6: Mean distance of occupational switches (conditional on switching) by age for different cohorts and education groups



high school or less switch to more different jobs as they age. The pattern for more educated workers is different: it exhibits a U-shape with age. Note that the rise in distance for later cohorts generally occurs throughout the entire life cycle.

B.2 Analysis of CPS Data

In this Appendix, I present a preliminary analysis of the evolution of the distance measure over time using CPS data matched with O*Net characteristics. I also describe some of the challenges presented by this dataset.

B.2.1 CPS Data

Description

The Current Population Survey (CPS), the source of many official Government statistics, is administered by the Bureau of the Census under the auspices of the Bureau of Labor Statistics (BLS). Currently, a nationally representative sample of about 65,000 households are interviewed monthly. Each household is interviewed once a month for four consecutive months one year, and again for the corresponding four month period a year later, resulting in 8 total months in the survey. Each month, new households are added and old ones that completed 8 months in the survey are dropped. Thus, eight rotation groups (cohorts of households starting their interviews in the same month) are interviewed in any given month (Kambourov and Manovskii, 2004).

Moscarini and Thomsson (2007) argue that the CPS is the only dataset that contains a fully representative sample of the US population, and only the CPS can potentially eliminate most sampling error when measuring employment distributions across small cells (e.g. 3-digit occupations, birth cohorts, etc.). The main disadvantage of the CPS for my purposes is that it is address-based, thus generating attrition when geographical mobility occurs, and geographical mobility may be correlated with occupational mobility. Also, the longitudinal dimension of the CPS is very limited. Another limitation of the CPS is that their coding of occupations changes with every decennial census.

Occupational transitions

The March Supplement of the CPS directly asks individuals about their current occupation, and the occupation they worked in during the previous year. It is available through IPUMS (King et al., 2010). Over time, the CPS has used several different occupational coding schemes (Census codes or 'COC'). 1968-70 was coded in 1960 COC; 1971-82 in 1970 COC; 1983-91 in 1980 COC; 1992-2002 in 1990 COC; and 2003-09 in 2000 COC. In any given year, both the individual's current occupation and previous year's occupation are classified in the same coding scheme. However, the changes in COC change the universe of occupations, making it impossible to make comparisons across COC-periods when using measures constructed on the basis of the occupational codes.[88] For example, the 1970 COC has 375

[88] An exception is the change between 1980 and 1990 COC which was relatively minor and therefore does not create an important break in the data

occupational codes, while the 1980 COC has 438 (Moscarini and Thomsson, 2007). The switch to a finer level of occupational coding would tend to mechanically increase measured occupational mobility, or, in my case, reduce mean distances (as similar occupations get split into different codes).[89]

Another limitation with the CPS has to do with changes in the imputation procedure. Missing records in the CPS are imputed through a 'hot deck' procedure, which seeks to match the individual with a missing record to another individual who is demographically similar. The missing data is imputed based on the answers provided by the demographically similar individual. The imputation procedure experienced significant changes in 1976 and 1988, leading to breaks in the series of occupational mobility in those years.

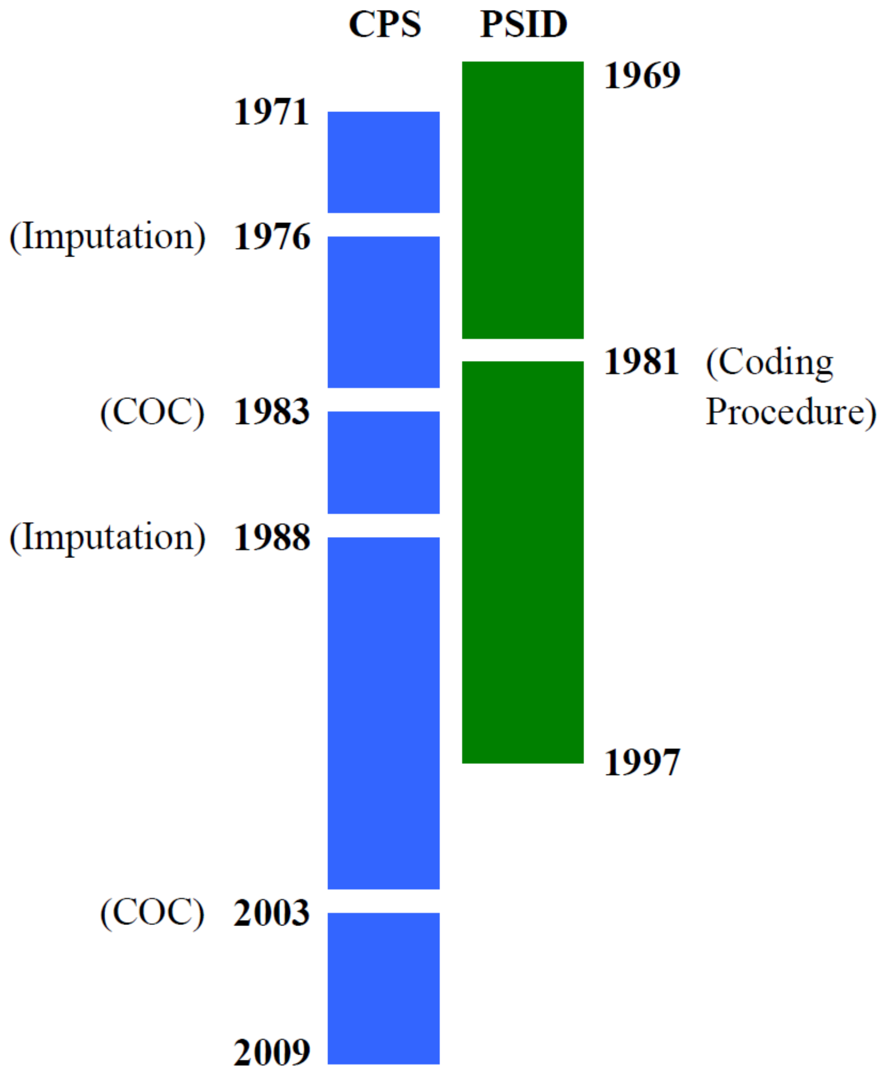
An advantage of the CPS occupational coding is that, starting in 1994, 'dependent-coding' techniques were introduced. This had the aim to reduce the interview burden and the possibility of misclassification. Instead of having individuals verbally describe their current and previous year's occupation, and coding each of them independently, individuals were directly asked whether they held the same job in both years. If they responded affirmatively, both the current and the previous year's occupation automatically received the same code. For more details on the depending coding procedure, and some potential problems in the way in which it is structured, see Moscarini and Thomsson (2007) and Kambourov and Manovskii (2004).

Because of the coding changes mentioned above, I can only analyze subperiods for which the coding was consistent, that is: 1971-75, 1976-82, 1983-87, 1988-2002, and 2003-09.[90] Figure B.7 indicates the sub-periods over which I have consistent data on occupational transitions from the CPS and the PSID, respectively. The reasons causing breaks in the data are also indicated. The overlapping periods for which there is consistent data in both the CPS and the PSID are 1971-75, 1976-80, 1983-87, and 1988-97.

[89] In terms of monthly occupational mobility, however, Moscarini and Vella (2008) find no evidence of major discontinuities upon reclassifications.

[90] An alternative option would be to convert all of the occupational codes to a common ('standardized') coding scheme, e.g. with a crosswalk like the one created by Peter Meyer and Anastasiya Osborne at the BLS (Meyer and Osborne, 2005). However, the disadvantage of doing this is that the matching across COCs is not perfect, so some of the standardized codes do not have a correspondence in all COCs. Also, it requires an important degree of aggregation of occupations. Moreover, even if all the codes are converted to a common code, the underlying universe of occupations is still changing, which still affects the measurements.

Figure B.7: Periods with consistent data on occupational transitions from the CPS and the PSID



Note: 'Imputation' indicates a change in the imputation procedure. 'COC' indicates a change in the coding scheme used. 'Coding Procedure' indicates a change in the way the occupational coding was performed. (See details in text.)

B.2.2 Matching DOT/O*Net with CPS

The aggregation process for DOT and O*Net is described in this subsection and summarized in Table B.1 for the DOT and Table B.2 for O*Net. The DOT-77 and DOT-91 each have their own coding schemes, which are much more disaggregated than the COC. To aggregate to the 1970-COC level, I follow Autor et al. (2003). I use the April 1971 CPS Monthly File, in which experts assigned individuals both with 1970-COC and DOT-77 codes. Using the CPS sampling weights, I calculate means of each DOT task measure at the 1970-COC occupation level. DOT task measures are missing for a small number of occupations. Occupational switches involving these occupations represent less than 0.5% of observed switches, and must be dropped from the sample.

Next, I attach the DOT task measures at the 1970-COC level to the March 1982 CPS. This is the last March CPS coded at the 1970-COC level. IPUMS provides the variable *occ*1990 (hereafter 1990m-COC), a consistent coding scheme based on a modified version of the 1980 and 1990 COCs. I use CPS sampling weights to obtain means of each DOT task measure at the 1990m-COC occupation level.

At this stage I have missing DOT task measures for a larger number of 1990m-level occupations. This is because many occupations in the 1990m coding scheme do not have a correspondence to a 1970-level occupation. I must drop around a fifth of observed occupational transitions when calculating distance measures based on the DOT for years coded at the 1990m-COC level. (See Table B.1).

For DOT-91, I use a crosswalk between DOT-91 and DOT-77 codes provided by David Autor. I follow the same procedure as in the previous paragraphs to calculate means of DOT-91 tasks by 1970-COC and 1990m-COC occupations.

O*Net-02 is coded in ONET-SOC 2000 Codes. The 2000-COC is also based on SOC codes. A crosswalk between SOC and 2000-COC codes is available from the National Crosswalk Service Center.[91] The conversion from ONET-SOC 2000 Codes to 2000-COC requires the combination of only a small number of occupations,[92] so I calculate simple averages for each O*Net dimension across the O*Net occupations that get combined into the same 2000-COC code. I have missing O*Net data for some 2000-COC

[91] `http://www.xwalkcenter.org/`

[92] The main exceptions are Engineering technicians (155 in 2000-COC), Postsecondary teachers (220), and Prepress technicians (825), which are subdivided into 10 or more detailed occupations in the SOC coding scheme.

occupations. This is because the corresponding O*Net codes have missing data. The vast majority of these occupations (weighted by the number of transitions to and from them) are 'not elsewhere classified' occupations, e.g. 'Managers, All Other', 'Sales and Related Workers, All Other', 'Production Workers, All Other'. I must drop these transitions which represent up to a fifth of observed transitions.

I then append the O*Net scores at the 2000-COC level to the March 2003 CPS. This is the first March CPS that has occupations coded at the 2000-COC level. As with the DOT, I use CPS weights to calculate means of each O*Net score by 1990m-COC occupation. Approximately a fifth of transitions occur between 1990m-COC occupations for which I have missing O*Net data (see Table B.2), in most cases because the 1990m-COC code does not have a corresponding 2000-COC code.

The whole aggregation processed is summarized graphically in Figure B.8. Tables B.1 and B.2 report a measure of the degree of similarity of the occupations getting combined upon each aggregation, 'Mean σ_{within} / $\sigma_{overall}$'. It is calculated as follows: I first calculate the standard deviation across the entire set of occupations for each DOT/O*Net dimension i (aptitudes/abilities), and call this $\sigma^i_{overall}$. This gives a sense of the dispersion of each dimension across the universe of occupations. Next, I calculate the standard deviation for each dimension i within each group of occupations that get combined when transitioning between coding schemes. I call this σ^i_{within}. I then calculate the mean of σ^i_{within} across all the sets of occupations and call this $\overline{\sigma}^i_{within}$. Finally, I calculate the ratio of $\overline{\sigma}^i_{within}$ and $\sigma^i_{overall}$ and average across all dimensions, as follows:

$$\frac{1}{I}\sum_{i=1}^{I}\frac{\overline{\sigma}^i_{within}}{\sigma^i_{overall}} \tag{B.2}$$

where I is the total number of dimensions in the DOT/O*Net being considered (as described in Section 3.2.4). The obtained values are reported in the Tables under 'Mean σ_{within} / $\sigma_{overall}$', and show that the occupations that get combined when going from one coding scheme to another are, in fact, similar to each other.

Each task measure in both the DOT and the O*Net is normalized so that the maximum (potential) score is equal to 10, and the minimum (potential) score is 0. Once I have DOT-77 and DOT-91 measures at the 1970-COC and the 1990m-COC levels, as well as O*Net-02 measures at the 1990m-COC and the 2000-COC levels, I can attach these to the corresponding data for each year of the March CPS. Specifically, DOT scores at the 1970-COC

Table B.1: Conversion of DOT data to different coding levels

Coding schemes	DOT → 1970-COC	1970-COC → 1990m-COC
Weights used	April 1971 CPS	March 1982 CPS
Gender specific	Yes	Yes
Universe of occupations	3,886 → 419	416 → 302
Mean σ_{within} / $\sigma_{overall}$	0.7220	0.3319
Attached to:	PSID: 1969-1997 CPS: 1971-1982	CPS: 1983-2002
Fraction of transitions with missing distance	PSID: 0.13% CPS: 0.28%	CPS: 20.64%

Note: DOT refers to the DOT-77 coding scheme.
For DOT-91, the codes are first converted to DOT-77 using a crosswalk provided by David Autor.

level can be attached to the 1971-1982 CPS, DOT and O*Net scores at the 1990m-COC level can be attached to the 1983-2002 CPS, and O*Net scores at the 2000-COC level can be attached to the 2003-2009 CPS.

B.2.3 Results

For the analysis using CPS data, I exclude two specific job transitions which experience unreasonable fluctuations in their frequencies over time. They are: 'Sales Representatives, Wholesale and Manufacturing' to 'Retail Salesperson', and 'Licensed Nurse' to 'Registered Nurse'. Figure B.9 presents median distance for occupational switchers in the CPS. Figure B.10 presents a similar graph for the sample restricted to male household heads working in the private sector. This dataset does not seem to show a clear trend in distance for occupational switchers.

Figure B.8: Graphic summary of procedure followed to aggregate DOT and O*Net codes to the different COC-code levels

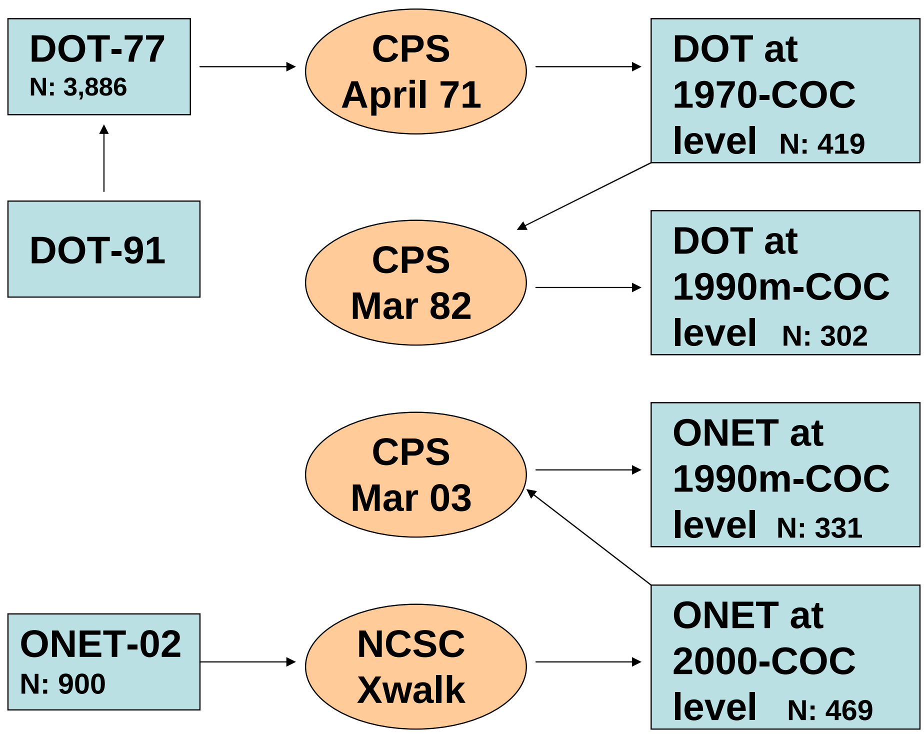


Table B.2: Conversion of O*Net data to different coding levels

Coding schemes	O*Net → 2000-COC	2000-COC → 1990m-COC
Weights used	None	March 2003 CPS
Gender specific	No	Yes
Universe of occupations	900 → 469	465 → 331
Mean σ_{within} / $\sigma_{overall}$	0.4848	0.4934
Attached to:	CPS: 2003-2009	CPS: 1983-2002
Fraction of transitions with missing distance	16.51%	11.70%

Note: O*Net refers to the ONET-SOC 2000 Codes.

Figure B.9: Median distance for occupational switchers, CPS, all household heads

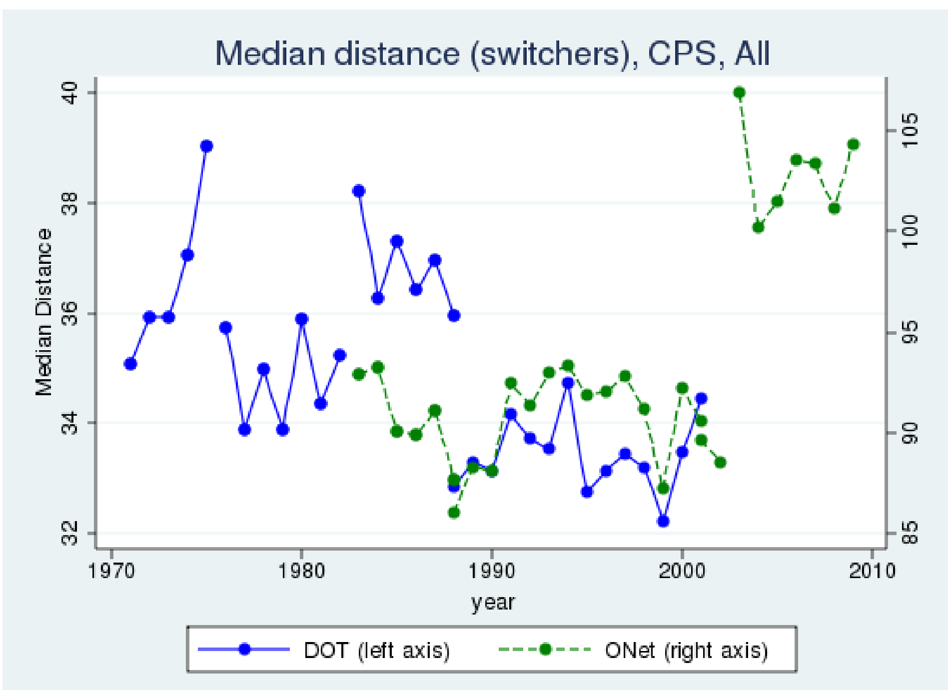
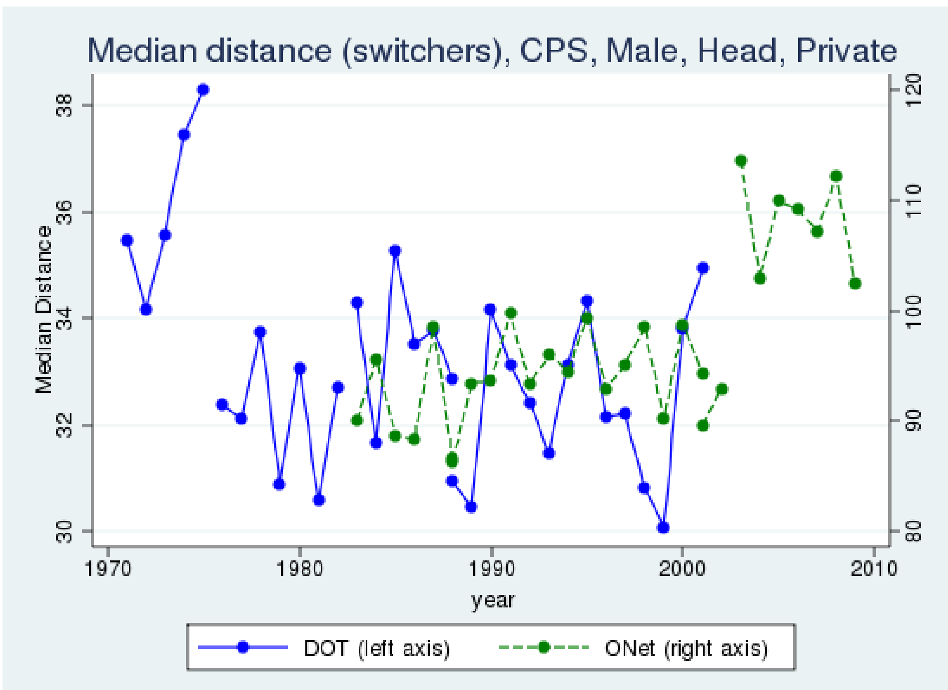


Figure B.10: Median distance for occupational switchers, CPS, only male private-sector household heads



B.3 Description of PSID Variable Construction

- Employer switches and tenure

Employer switches are constructed following 'Partition T' described in Kambourov and Manovskii (2009b). Based on these employer switches, a consistent series of employer tenure is constructed.

- Occupation and industry switches

An occupation (industry) switch is recorded when a person's current occupation (industry) report is different from his or her occupation (industry) report in the previous calendar year. For the period plagued with coding error (1981 onwards - see main text for details on the change in coding procedure), I sometimes report results for 'restricted' occupation (industry) switches. These are cases where the individual's occupation (industry) switch is accompanied by an employer switch (see Kambourov and Manovskii (2009b) for details on why this is a reasonable restriction).

- Occupation and industry tenure

These variables are calculated following the 'Employer-T' procedure described in Kambourov and Manovskii (2009b). Basically, for the purpose of building occupation and industry tenure, an occupation or industry switch is recorded only if an employer switch is recorded using 'Partition T'.

- Wages

Wages are reported by the individuals for their main job at the time of the interview, and can therefore be related to their current occupation. For the case of workers who report a salary (instead of an hourly wage), an hourly wage equivalent is calculated. Consistent with PSID procedures, if the worker reports a weekly salary, it is divided by 40 hours; a bi-weekly salary is divided by 80; a monthly salary by 160; and an annual salary by 2000. Wages are available from 1971 onwards. For salaried workers, they are only available from 1976 onwards for household heads, and from 1979 onwards for wives. Hourly wages are top-coded at $9.98 up to 1977. All nominal variables are converted to real 1979 dollars using the Bureau of Labor Statistic's Consumer Price Index (CPI) for all urban consumers and all items.

- Education

Each individual's education level is fixed for all years at their highest reported education level in any year in which they are in the sample.

- Marital status

Cohabitors are treated as married up to 1976. After that, only legally married couples are treated as married.

- SMSA

The variable takes a value of 1 if the largest city in the respondent's county of residence has a population of at least 50,000 people.

- Age

Age is adjusted so that it increases by 1 every year.

Appendix C

Appendices to Chapter 4

Table C.1: 2-digit occupation groupings for the Meyer and Osborne (2005) coding system

2-digit Category	2-digit Code	3-digit 'Meyer-Osborne' Code
Public administrators and officials	01	003-004
Other executive, administrators, and managers	02	007-022
Management related occupations	03	023-037
Engineers, architects	04	043-059
Mathematical and computer scientists	05	064-068
Natural scientists	06	069-083
Health diagnosing occupations	07	084-089
Health assessment and treating occupations	08	095-106
Teachers, college and university	09	113-154
Teachers, except college and university	10	155-163
Librarians, social scientists, religious workers	11	164-176
Lawyers and judges	12	178-179
Writers, artists, entertainers, athletes	13	183-199
Health technologists and technicians	14	203-208
Engineering and science technicians	15	213-225
Technicians, except health engineering, and science	16	226-235
Sales supervisors and sales reps, finance and business	17	243-256
Retail and other salespersons	18	258-283
Office supervisors and computer operators	19	303-308
Secretaries, stenographers, and typists	20	313-315
Information and records processing, except financial	21	316-336
Financial records processing occupations	22	337-344
Office machine and communication equipment operators	23	345-349
Mail and message distributing	24	354-357
Other administrative support occupations, including clerical	25	359-389

Table C.1: 2-digit occupation groupings for the Meyer and Osborne (2005) coding system

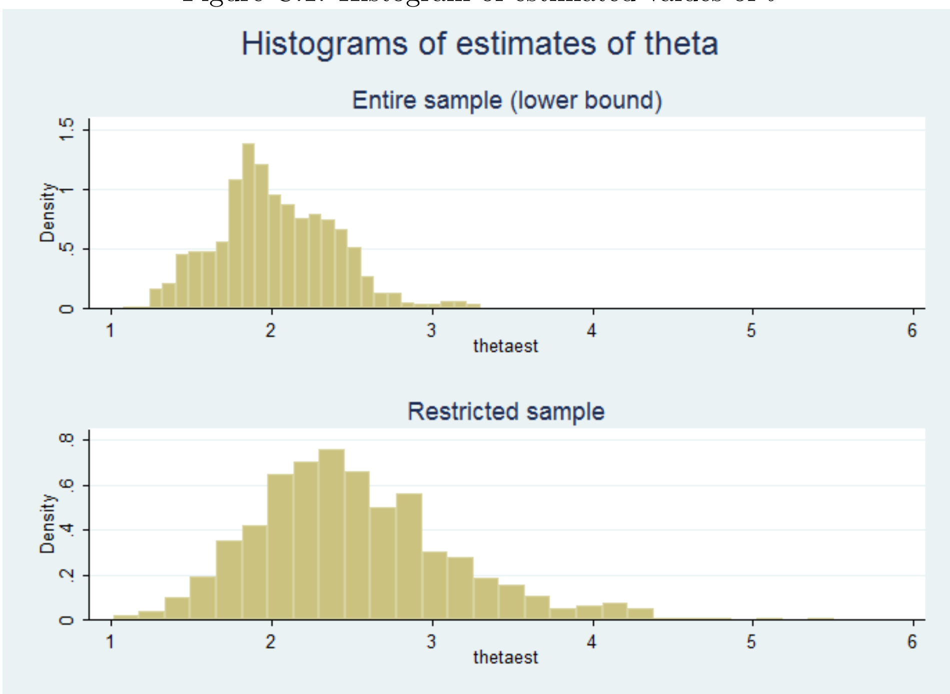
Private household cleaners and servers	26	405-407
Protective service occupations	27	415-427
Food service occupations	28	434-444
Health service occupations	29	445-447
Cleaning and building service occupations, except household	30	448-455
Other personal service occupations	31	456-469
Farm operators and managers	32	473-476
Farm workers and related occupations	33	479-489
Forestry and fishing occupations	34	496-498
Mechanics and repairers	35	503-549
Construction trades	36	558-599
Other precision production occupations	37	614-699
Machine operators and tenders, not precision	38	703-779
Fabricators, assemblers and hand working occupations	39	783-789
Production inspectors and graders	40	796-799
Motor vehicle operators	41	803-813
Other transportation and material moving	42	823-859
Helpers, construction and production occupations	43	865-874
Freight, stock and material handlers	44	875-889
Military occupations, unemployed or unknown	45	905-999

C.1 Matching DOT with CPS

The National Crosswalk Service Center provides a crosswalk between the occupation codes in the 1991 Dictionary of Occupational Titles (DOT) and the 1990 Census Occupation Codes (COC).[93] 1990-COC codes are first converted to Meyer and Osborne (2005) standardized codes. Next, because the DOT classification is much more detailed than the 1990-COC, unweighted means for all the DOT codes that fall within the same standardized code are calculated. Each dimension of the DOT is then rescaled to have mean zero and standard deviation of one across the universe of (unweighted) standardized Meyer-Osborne occupations. Finally, to generate scores at the 2-digit

[93] The crosswalk is the National Occupational Information Coordination Committee (NOICC) Master Crosswalk, Version 4.3, downloadable from `ftp://ftp.xwalkcenter.org/download/xwalks/`, file `xwalkv43.exe`.

Figure C.1: Histogram of estimated values of θ



Note: The histogram for the entire sample includes estimated values of θ for each occupation with at least 100 reports. The histogram for the restricted sample includes estimated values of θ for each education group by initial occupation, restricted to workers aged 25-30, and to groups with at least 20 earnings reports. Estimates are annual over the period 1983-2002.

Meyer-Osborne level, an unweighted average is taken across all 3-digit occupations that are within the same 2-digit category.

C.2 Extension: Occupation Tenure

This section extends the model to allow for occupation-specific human capital.[94] Let an individual's tenure in occupation j be denoted $ten_j(i)$, and assume that occupational tenure increases productivity at a rate of γ for each additional year of tenure. This leads to the following modified version of Equation (4.1):

$$w_j(i) = p_j f\left[X(i)\right](1 + ten_j(i))^{\gamma}\left(\frac{z_j(i)}{d_{kj}}\right) \tag{C.1}$$

The extra productivity from tenure is occupation-specific human capital. It is entirely non-transferable and lost when switching out of occupation j.[95]

With this modified wage specification, the probability that occupation j offers individual i the highest wage, which is the probability that individual i will optimally choose to switch to occupation j, given his current occupation k (denoted by $\pi_{kj}(i)$) is now given by:

$$\begin{aligned}\pi_{kj}(i) &\equiv Pr\left[w_j(i) \geq \max_s \{w_s(i)\}\right] \\ &= \int_0^{\infty} Pr\left[w_s(i) \leq w, \forall s \neq j\right] \cdot dPr\left[w_j(i) \leq w\right] \\ &= \frac{T_j d_{kj}^{-\theta}\left[p_j\left(1 + ten_j(i)\right)^{\gamma}\right]^{\theta}}{\sum_{s=1}^{N} T_s d_{ks}^{-\theta}\left[p_s\left(1 + ten_s(i)\right)^{\gamma}\right]^{\theta}}\end{aligned} \tag{C.2}$$

Note that $ten_j(i) = 0 \ \forall j \neq k$. Therefore, $\forall j \neq k$:

$$\pi_{kj}(i) = \frac{T_j d_{kj}^{-\theta} p_j^{\theta}}{\sum_{s \neq k} T_s d_{ks}^{-\theta} p_s^{\theta} + T_k p_k^{\theta}\left(1 + ten_k(i)\right)^{\gamma\theta}} \tag{C.3}$$

[94] For evidence on the importance of occupation-specific human capital, see Kambourov and Manovskii (2009b).

[95] I assume that tenure is reset to zero when an individual switches out of an occupation. Therefore, if an individual switches back to an occupation after working in a different occupation for a certain number of years, he restarts with zero tenure. (The same assumption is made by Kambourov and Manovskii (2009b).) Note that occupation-specific human capital is transferable across employers within the same occupation.

Meanwhile, individual i's probability of staying in occupation k, π_{kk} is given by:

$$\pi_{kk}(i) = \frac{T_k p_k^\theta \left(1 + ten_k(i)\right)^{\gamma\theta}}{\sum_{s \neq k} T_s d_{ks}^{-\theta} p_s^\theta + T_k p_k^\theta \left(1 + ten_k(i)\right)^{\gamma\theta}} \tag{C.4}$$

Dividing (C.3) by (C.4), and taking logs of the ratio, we have:

$$\ln \frac{\pi_{kj}(i)}{\pi_{kk}(i)} = \ln T_j + \theta \ln p_j - \ln T_k - \theta \ln p_k - \theta \ln d_{kj} - \gamma\theta \ln(1 + ten_k(i)) \tag{C.5}$$

Averaging this across individuals in occupation k leads to the gravity-type equation:

$$\begin{aligned} \frac{1}{N_k} \sum_{i=1}^{N_k} \ln \frac{\pi_{kj}(i)}{\pi_{kk}(i)} &= \ln T_j + \theta \ln p_j - \ln T_k - \theta \ln p_k \\ &\qquad - \theta \ln d_{kj} - \gamma\theta \frac{1}{N_k} \sum_{i=1}^{N_k} \ln(1 + ten_k(i)) \end{aligned} \tag{C.6}$$

where N_k is the number of individuals in occupation k.

In order to estimate Equation (C.6) in the data, I need an empirical counterpart of the term on the left-hand-side of the equation. It is an average of individual-specific probability ratios. Note, however, from (C.3) and (C.4), that the only individual-specific component in $\pi_{kj}(i)$ and $\pi_{kk}(i)$ is occupational tenure. All individuals with tenure level x have the same transition probabilities.

Let π_{kj}^x be the probability of switching between k and j for individuals with tenure level x. I can rewrite the left-hand-side of equation (C.6) as a weighted average across tenure groups in the following way:

$$\frac{1}{N_k} \sum_{i=1}^{N_k} \ln \frac{\pi_{kj}(i)}{\pi_{kk}(i)} = \frac{1}{N_k} \sum_x N_k^x \ln \frac{\pi_{kj}^x}{\pi_{kk}^x} \tag{C.7}$$

where N_k^x is the number of individuals in occupation k with tenure level x (at the start of the period), and the sum is over the different levels of x.

Assume that there is a large number of individuals in each occupation and at each tenure level. The probability of switching between k and j (or staying in k) given a tenure level of x is therefore equal to the fraction of workers of tenure level x in occupation k who optimally choose to switch to j (or stay in k). That is:

$$\pi^x_{kj} = \frac{sw^x_{kj}}{N^x_k} \tag{C.8}$$

$$\pi^x_{kk} = \frac{sw^x_{kk}}{N^x_k} \tag{C.9}$$

where sw^x_{kj} stands for the number of switchers from occupation k to occupation j with tenure x.

Therefore, substituting (C.8) and (C.9) into (C.7), implies that the empirical counterpart of the left-hand-side of equation (C.6) is given by:

$$\frac{1}{N_k}\sum_{i=1}^{N_k} \ln \frac{\pi_{kj}(i)}{\pi_{kk}(i)} = \frac{1}{N_k}\sum_{x} N^x_k \ln \frac{sw^x_{kj}}{sw^x_{kk}} \tag{C.10}$$

Here I encounter a data limitation. Measurement of the right-hand-side of equation (C.10) requires a large dataset with information on flows of workers across particular occupation pairs, as well as tenure data. An ideal dataset to measure flows of workers across occupation pairs, due to its large sample size, is the CPS. However, this dataset has little information on occupational tenure. On the other hand, the PSID provides information that allows the construction of measures of occupational tenure.[96] However, the sample size in the PSID is much smaller and does not reliably capture flows of workers across particular occupation pairs for different tenure levels.

I show now how I deal with this limitation by combining information from both datasets for the empirical measurement of the left-hand-side of equation (C.6).

The CPS provides data on sw_{kj}/sw_{kk}. It can be shown that, when there

[96] See Kambourov and Manovskii (2009b).

is a large number of individuals in each occupation:[97]

$$\ln\left(\frac{\sum_{i=1}^{N_k}\pi_{kj}(i)}{\sum_{i=1}^{N_k}\pi_{kk}(i)}\right) = \ln\left(\frac{sw_{kj}}{sw_{kk}}\right) \tag{C.11}$$

It can also be shown that the difference between the object of interest (equation (C.10)) and the data from the CPS (equation (C.11)) is constant for each k (source occupation). That is:[98]

$$\ln\left(\frac{\sum_{i=1}^{N_k}\pi_{kj}(i)}{\sum_{i=1}^{N_k}\pi_{kk}(i)}\right) - \frac{1}{N_k}\sum_{i=1}^{N_k}\ln\frac{\pi_{kj}(i)}{\pi_{kk}(i)} = constant_k \tag{C.12}$$

Therefore, the bias between $\frac{1}{N_k}\sum_x N_k^x \ln\frac{sw_{kj}^x}{sw_{kk}^x}$ and $\ln(sw_{kj}/sw_{kk})$ is constant by source occupation. If I use the latter as the dependent variable, this measurement bias can be captured with a source-occupation fixed effect. Alternatively, one can get a direct measure of the bias using PSID data.

[97] Following the same logic from Equations (C.7) through (C.10), with a large number of individuals in each occupation and at each tenure level:

$$\begin{aligned}
\frac{1}{N_k}\sum_{i=1}^{N_k}\pi_{kj}(i) &= \frac{1}{N_k}\sum_x N_k^x\pi_k^x \\
&= \frac{1}{N_k}\sum_x N_k^x\left(\frac{sw_k^x}{N_k^x}\right) \\
&= \frac{sw_{kj}}{N_k}
\end{aligned}$$

[98] This comes from the fact that:

$$\begin{aligned}
&\ln\left(\frac{\sum_{i=1}^{N_k}\pi_{kj}(i)}{\sum_{i=1}^{N_k}\pi_{kk}(i)}\right) - \frac{1}{N_k}\sum_{i=1}^{N_k}\ln\frac{\pi_{kj}(i)}{\pi_{kk}(i)} \\
=&\ln\left(\frac{\sum_{i=1}^{N_k}\left(T_j d_{kj}^{-\theta}/denom_k(i)\right)}{\sum_{i=1}^{N_k}\pi_{kk}(i)}\right) - \frac{1}{N_k}\sum_{i=1}^{N_k}\ln\frac{\left(T_j d_{kj}^{-\theta}/denom_k(i)\right)}{\pi_{kk}(i)} \\
=&\ln\left(T_j d_{kj}^{-\theta}\right) + \ln\left(\frac{\sum_{i=1}^{N_k}(1/denom_k(i))}{\sum_{i=1}^{N_k}\pi_{kk}(i)}\right) - \ln\left(T_j d_{kj}^{-\theta}\right) - \frac{1}{N_k}\sum_{i=1}^{N_k}\ln\frac{(1/denom_k(i))}{\pi_{kk}(i)} \\
=&\ln\left(\frac{\sum_{i=1}^{N_k}(1/denom_k(i))}{\sum_{i=1}^{N_k}\pi_{kk}(i)}\right) - \frac{1}{N_k}\sum_{i=1}^{N_k}\ln\frac{(1/denom_k(i))}{\pi_{kk}(i)} \\
=&constant_k
\end{aligned}$$

where I have used the definition of π_{kj} from Equation (C.3), renaming the denominator as $denom_k(i)$, and the last line comes from the fact that $\pi_{kk}(i)$ and $denom_k(i)$ are independent of j.

In particular, equation (C.12) applies not only to π_{kj}, but also to the total flow of switchers ($\sum_{j\neq k}\pi_{kj}$). From the PSID I can calculate the ratios of total switchers over total stayers, by tenure level and by initial occupation. With this, I can measure:

$$\ln\frac{switchers_k}{stayers_k}-\frac{1}{N_k}\sum_x N_k^x \ln\frac{switchers_k^x}{stayers_k^x}=constant_k \tag{C.13}$$

The adjustment factor $constant_k$ can then be used to convert all measures of $ln(sw_{kj}/sw_{kk})$ to the object of interest.

Using Equations (C.11) and (C.12), the gravity equation (4.11) can be rewritten as:

$$\begin{aligned}\ln\frac{sw_{kj}}{sw_{kk}} &= \ln T_j+\theta\ln p_j-\ln T_k-\theta\ln p_k+constant_k\\ &\quad -\theta\ln d_{kj}-\gamma\theta\frac{1}{N_k}\sum_{i=1}^{N_k}\ln(1+ten_k(i))\end{aligned} \tag{C.14}$$

Define $D_j \equiv \ln T_j+\theta\ln p_j$, and $S_k\equiv D_k-constant_k$. The equation therefore becomes:

$$\ln\frac{sw_{kj}}{sw_{kk}}=D_j-S_k-\theta\ln d_{kj}-\gamma\theta\frac{1}{N_k}\sum_{i=1}^{N_k}\ln(1+ten_k(i)) \tag{C.15}$$

This suggests running the same equation as in the main body of the paper, but adding the average log-tenure for the source occupation.[99] I can use data on occupational tenure from the PSID to measure average tenure by source occupation.

Results are in Table C.2. Tenure does not seem to play an important role. Note that the effect of tenure is identified off of variation over time within each occupation, because the estimating equation includes source occupation fixed effect. As there is little variation over time in average tenure within an occupation, adding tenure does not affect the results.

[99] This would be captured by the fixed effect, but this particular variable can be calculated for each period t.

Table C.2: Estimated coefficients on 'gravity-type' equation including mean tenure as a control, 1983-2002

	OLS All	OLS Non-zero	Tobit
	(1)	(2)	(3)
logd91	-.858	-.398	-1.331
	(.020)***	(.008)***	(.034)***
common-maj	1.039	.506	1.580
	(.041)***	(.016)***	(.068)***
common-tint	.245	.041	.467
	(.031)***	(.013)***	(.053)***
(mean) logoccten	.087	.040	.076
	(.110)	(.057)	(.204)
Obs.	35131	19709	35131
R^2	.451	.592	

Note: Dependent variable is $\ln(sw_{kj}/sw_{kk})$. Regression includes source and destination occupation dummies, as well as year dummies.