Title
Essays on the Role of the Firm in Labor Economics

Author
Smith, Benjamin Simpson

Publication Date
2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Essays on the Role of the Firm in Labor Economics

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Economics

by

Benjamin Simpson Smith

2018

ABSTRACT OF THE DISSERTATION

Essays on the Role of the Firm in Labor Economics

by

Benjamin Simpson Smith

Doctor of Philosophy in Economics

University of California, Los Angeles, 2018

Professor Till von Wachter, Chair

The first chapter of this dissertation studies the causes of rising sorting between high-skill workers and high-paying firms. Despite accounting for a substantial share of rising wage inequality, little is known about how or why sorting is rising. To understand how, I develop a novel decomposition method to measure the relative importance of different worker flow channels. I find that labor market entry of young workers accounts for about half of the total rise in sorting. To understand why sorting is rising, I use exogenous variation induced by the fall of the Soviet Union to estimate the effect of trade liberalization on rising sorting within German local labor markets. I find that export exposure can account for 14% of the rise in sorting. I then apply the decomposition method to the export-induced changes in employment to confirm an important role for labor market entry in rising sorting.

The second chapter studies the effect of temporary employment shocks on the future earnings of professional golfers. Although a large literature documents the persistent effects of temporary employment shocks on the earnings of wage-and-salary workers, we have little evidence on the effects on self-employed workers. I exploit entry rules of the PGA TOUR to estimate the long-

term effects of temporary employment shocks using a regression discontinuity design. Although, I find large earnings differences in first year after an employment shock, these differences quickly dissipate. Furthermore, I find no effects of employment shocks on performance. Golfers have less job stability than typical workers and these differences likely explain why the earnings losses of golfers are less persistent than of wage-and-salary workers.

The third chapter studies the evolution of wages at large firms. Although large firms have paid significantly higher wages for over a century, we document that the large-firm wage premium has declined over the last thirty years. Decomposing pay into worker and firm fixed effects, we show that the decline is due to a reduction in firm effects at large firms, while worker composition has changed little. We also find that the majority of the change occurs within industries.

The dissertation of Benjamin Simpson Smith is approved.

Pablo David Fajgelbaum

Adriana Lleras-Muney

Nico Voigtlaender

Till von Wachter, Committee Chair

University of California, Los Angeles

2018

To all of the teachers who encouraged and inspired me, including: my gradmother, Maryan Smith, who said to "aim your arrow high," many at Ruffing Montessori, Dennis Arko at St. Ignatius High School, Mike Curme and Rich Hart at Miami University, and Monica Visan at the University of Chicago.

Contents

List of Figures

List of Tables

Acknowledgments

I am very grateful to my advisor, Till von Wachter, from whom I learned much about labor economics and about how to conduct research. Thanks for the discussions about economics and research ideas, for working so hard to find me great data, and for the stimulating collaborations over the years. I'd also like to thank the other members of my committee, Adriana and Pablo, for their suggestions, advice, and support on the job market. I am also grateful for the support and assistance of Nick Bloom and Fatih Guvenen who both went out of their way to help improve my research despite the fact that I was a student at a different university.

I am grateful to the California Center for Population Research (CCPR) for hosting me for the majority of my graduate school experience. In particular, I'd like to thank Ana Ramirez for outstanding administrative support. It has been an excellent experience and I am proud to be a member of the CCPR community. I'm also grateful to the UCLA Department of Economics including Jessica Perez, Chiara Paz, and Juliana Smith, for handling my various requests over the years and doing their best to make my job market experience go smoothly. I am grateful to the students and faculty involved in the applied seminars at UCLA for many helpful suggestions, for helping me develop the ability to present research, and for fostering an excellent research environment. Thanks also to the members of my second and third year research group, Owen Hearey, Matt Miller, and Chad Stecher, for camaraderie and direction.

I am particularly grateful to my colleagues and friends Ioannis Kospentaris and Richard Domurat. They have sat through many of my presentations and read many of my drafts. Thank you for all your time, effort, encouragement, and extremely helpful suggestions. Thanks for all the interesting conversations about research, economics, and philosophy. Most importantly, thanks for being such excellent friends and comrades during this voyage.

Most importantly, I am grateful to Andy for her patience, guidance, and encouragement throughout this process. Her ceaselessly positive attitude is, in large part, responsible for my completion of this dissertation. Finally, I'd like to thank my parents, Bob and Susie, and brothers, Nick and Ted, for their unwavering support and understanding of this strange career choice that took a decade to come to fruition.

Chapter 1 was produced with data from the German Social Security system. I'd like to thank The Research Data Centre (FDZ) of the Germany Federal Employment Agency (BA) at the Institute for Employment Research (IAB) for access to the data and for handling my many questions and requests.

Chapter 2 was produced with data from professional golf tours from around the world and, in particular, the PGA TOUR. I'd like to thank the staff of the PGA TOUR including Kristen Burgess, Dana Davies, and Ken Lovell for their patience and diligence in responding to my many requests. I also thank the PGA TOUR for providing access to ShotLink™data.

Powered By SHOTLink

Chapter 3 is a version of Bloom, Nicholas, Fatih Guvenen, Benjamin S. Smith, Jae Song, and Till von Wachter. 2018. "The Disappearing Large-Firm Wage Premium." *AEA Papers and Proceedings*, 108 : 317-22. Copyright American Economic Association; reproduced with permission of American Economic Association Papers and Proceedings.

Biography

Prior Education

Master of Arts in Economics, University of California, Los Angeles 2014

Graduate Student-at-Large, University of Chicago 2008 – 2011

B.S.B. in Business Economics and Finance, Miami University 2007

Publications

The Disappearing Large Firm Wage Premium
(with Nicholas Bloom, Fatih Guvenen, Jae Song, and Till von Wachter)
AEA Papers and Proceedings, May 2018

Working Papers

The Role of Labor Market Entry and Trade in Sorting: Evidence from West Germany

The Effects of Employment Shocks on the Self-Employed: Evidence from Natural Experiments in Professional Golf

Earnings Growth, Job Mobility, and Inequality from 1957 to 2014
(with Jae Song and Till von Wachter)

Invited Presentations

Federal Trade Commission, Analysis Group, Charles River Associates	2018
Federal Reserve Bank of Minneapolis	2017
Society for Institutional & Organizational Economics	2017

Teaching Experience

Teaching Assistant, Microeconomic Theory I	UCLA, Fall 2014
Teaching Assistant, Introduction to Econometrics	UCLA, Spring 2014
Teaching Assistant, Introduction to Econometrics	UCLA, Winter 2014
Teaching Assistant, Microeconomic Theory I	UCLA, Fall 2013

Research Assistance

Research Assistant, Till von Wachter, University of California, Los Angeles	2014 – 2017
Summer Research Internship, United States Social Security Administration	2015
Research Assistant, Matthew Notowidigdo, University of Chicago	2011 – 2012
Research Assistant, Precision Health Economics	2010 – 2011
Research Assistant, Becker Center on Chicago Price Theory	2009 – 2010

Honors, Awards, Scholarships, & Professional Activites

Dissertation Year Fellowship, UCLA	2017 – 2018
Graduate Research Mentorship, UCLA	2015 – 2016
Price Theory Summer Camp, University of Chicago	2015
Summer School on Socioeconomic Inequality, University of Chicago	2015
Student Affiliate, California Center for Population Research	2014 – 2018
President, UCLA Graduate Economics Association	2016 – 2017

Chapter 1

The Role of Labor Market Entry and Exports in Sorting: Evidence from West Germany

1.1 Introduction

The rise in earnings inequality has been widely documented in many developed economies including Germany, the US, and the UK (e.g. Dustmann et al. 2009; Autor et al. 2008; Katz et al. 1993). Traditionally, economists have sought explanations for rising inequality in changing returns to worker characteristics, such as education, occupation, and skill (e.g. Katz and Autor 1999; Acemoglu and Autor 2011). However, recent studies find that earnings are also heavily determined by firm characteristics. This literature documents large variation in firm-specific wage premiums and, over time, an increasing employment of high earning individuals at high-wage firms.[1] In fact, this sorting amongst workers and firms accounts for about 30% of the rise in inequality in both Germany and the US (Card et al., 2013; Song et al., 2018).[2] Yet, despite accounting for a siz-

[1] For an account of the magnitude of variation due to firm wage premiums see, among others, Abowd et al. (1999); Goux and Maurin (1999); Abowd et al. (2002); Gruetter and Lalive (2009); Holzer et al. (2011); Song et al. (2018).

[2] Bagger et al. (2013) also find increased sorting in Denmark. Håkanson et al. (2015) find that firms are increasingly segregated in terms of worker skills in Sweden–a related concept to sorting.

able share of rising inequality, very little is known about how or why sorting is rising. To make progress towards this end, I perform three exercises. First, to understand how sorting is rising, I develop a novel decomposition method to quantify the importance of different worker flow channels. Second, to understand why sorting is rising, I estimate the causal effect of liberalization of international trade. Third, I combine the previous exercises by applying the worker flow decomposition to only the changes in employment induced by exogenous trade variation. Holding worker composition constant, this decomposition documents the channels by which labor demand leads to sorting.

I measure labor market sorting based on worker and firm wage components estimated in a fixed effects regression following the methodology of Abowd et al. (1999). The worker component represents the portion of earnings capacity which is portable across all jobs.[3] The firm component represents a firm-specific earnings premium paid to all employees regardless of worker ability.[4] Sorting is then defined as the correlation between worker and firm wage components across jobs. A positive sorting measure means that, on average, high-wage workers are employed at higher-wage firms than low-wage workers. Within this framework, even if the variance of both worker and firm wage components is constant over time, increased sorting can cause higher inequality. For example, given only two workers and two firms, wage variation will be greater if the lower-wage worker is employed by the lower-wage firm, rather than if the lower-wage worker is employed by the higher-wage firm.

Although researchers have identified an increase in sorting, how this process has taken place remains an open question. In particular, we lack an understanding of which worker flows have led to the rise. Sorting is often thought to arise through job-to-job transitions (e.g. Hagedorn et

[3]This component captures any fixed worker characteristics that affect earnings such as education or skill.

[4]One prominent explanation for the existence of firm wage components is that more successful firms share some portion of their rents with their employees. See Card et al. (2018) for an overview of the connection between rent-sharing and firm wage premiums.

al. 2017; Lopes de Melo 2018) since they account for a large fraction of employment separations (e.g. Fallick and Fleischman 2004) and help to match the magnitude of frictional wage dispersion in search models (Hornstein et al., 2011). However, we lack empirical evidence to verify this conjecture against alternative channels.

To bridge this gap, I develop a novel decomposition method to quantify the relative importance of different worker flow channels to the rise in sorting over time. I classify workers into the following worker flow groups according to their employment history: job-to-job transitions, labor market entry of young workers, and employment-to-nonemployment transitions.[5] By constructing counterfactual joint distributions by worker and firm wage components in which a given channel is held constant, I am able to compare the relative contributions of very different types of worker flows. As a result, this framework offers a comprehensive view of the factors that shape aggregate sorting patterns.

The results of this decomposition show that labor market entry is the most important worker flow leading to rising sorting. In fact, I estimate that labor market entry accounts for 57.0% of the rise. On the other hand, job-to-job transitions play a more limited role, accounting for only 17.8% (with an upper bound of 26.7%). The remaining fraction is due to job stayers (12.8%) and unemployment transitions (2.6%). This decomposition, which I label the *aggregate decomposition*, offers some of the first evidence as to how the sorting process occurs. Despite the focus on job-to-job transitions in the literature, increases in the initial sorting of young workers in their first jobs is a more important channel for rising sorting.

Next I move on to the question of why sorting is rising. Of the multiple potential causes, I focus on the role of trade.[6] Most theories of sorting are based on the idea that high-skill workers

[5]Employment-to-nonemployment transitions include transitions to unemployment and other unidentified states such as non-participation, self-employment, part-time employment, and employment in East Germany or foreign countries.

[6]Some other potential causes are technological change, changes in the skill distribution, and changes in the degree

are especially productive at high-productivity firms. By increasing the scale of the market, export exposure may strengthen these complementarities between workers and firms and lead to greater sorting (Bombardini et al., 2015). In addition to the theoretical interest in the effect of trade on sorting, trade liberalization is also an empirically significant channel. In fact, the value of exports as a percentage of German GDP rose from 22% in 1988 to 39% in 2006.[7]

To identify this channel I exploit the surge in German trade after the fall of the Soviet Union and the rise of China in the 1990s and 2000s. Following the methodology of Autor et al. (2013) and Dauth et al. (2014), I construct measures of import and export exposure based on the value of trade between Germany and Eastern Europe and China (the "East") at the local labor market level. The rise of trade with the East had differential impacts across German *industries* and, therefore, differential impacts across German *regions*. This variation can be credibly argued to be exogenous to domestic supply and demand variation across industries since these events were largely motivated by internal politics within the East. Furthermore, both regions ascended to World Trade Organization membership around 2001, representing a second exogenous shock to terms of trade. This research design aims to identify changes to labor demand induced by changes in foreign product supply and demand, while excluding effects due to changes in domestic factors.

I estimate a significant, positive causal effect of export exposure on local labor market sorting. In contrast, import exposure shocks have an insignificant, negative effect on sorting. Using the average change in trade exposure over the period, I find that trade shocks from the East can account for a substantial share (14%) of the total rise in West German sorting. Furthermore, I show that export exposure shocks increase manufacturing employment and wages and, therefore, substantiate the interpretation of export shocks as labor demand shocks. These results suggest that trade

of search frictions.

[7] As this statistic fluctuates annually, these numbers represent seven year averages in which the indicated year is the median of the interval (source: OECD).

liberalization is an important factor driving increased sorting.

In the final part of the paper, I connect the previous exercises to examine whether the increase in sorting at labor market entry is caused by changes in labor demand or labor supply. On the supply side, the aggregate decomposition may be affected by trends toward higher education at the top end and waves of low-skill immigration at the bottom end. In order to disentangle the effects of changes in the composition of supply from changes in demand, I apply the worker flow decomposition to changes in employment induced by exogenous variation in export exposure. Recall that the trade instrument is specifically designed to isolate the effect of demand. I then compare the aggregate decomposition with the export decomposition to understand whether the effects of demand are similar to the aggregate effects.

In performing this decomposition, which I label the *export decomposition*, I find very similar results to the aggregate decomposition. Labor market entry is the most important channel for rising sorting–accounting for 47.0%. Job-to-job transitions again account of limited share at 16.6% with an upper bound of 26.4%. Furthermore, I find that these entry effects are in large part driven by the entry of low-wage workers to low-wage service firms. Given that these sorting effects are driven by the demand side, the results suggest that, over time, even if the distribution of workers entering the labor market is the same, sorting will increase due to the demand effects of trade liberalization and other factors that work through similar channels.

1.1.1 Contribution to the Literature

My main contribution is to provide empirical evidence about the relative roles of different workers flows in labor market sorting. To date, the literature focuses on changes in the allocation of workers to firms after entry into the labor market, arising through job-to-job or employment-to-nonemployment (or unemployment) transitions. Most of the emphasis is on job-to-job transitions since they occur at about twice the rate of employment to unemployment transitions (Fallick and

Fleischman, 2004; Nagypál, 2008) and since they fuel the faster growth of high productivity firms (Haltiwanger et al., 2017a). As a result, models of sorting often incorporate on-the-job search in order to match data on worker flows.[8] These models are consistent with the idea that technological shocks require a reallocation of the workforce, which is accommodated through job-to-job transitions. To my knowledge, Haltiwanger et al. (2017b) provide the only direct empirical evidence to assess the role of job-to-job transitions in labor market sorting. Perhaps surprisingly, they find that job-to-job transitions act to mitigate assortative matching as low-wage, low-educated workers are more likely to move up the job ladder to high-wage, high-productivity firms. However, they do not produce a comprehensive framework to compare the effect of job-to-job transitions with other worker flows. In developing such a comprehensive framework, I find that one of the most important channels driving sorting, labor market entry, has heretofore been neglected by the literature.

The finding that the growth in sorting occurs at labor market entry has implications for the study of inequality and persistence in the labor market. Guvenen et al. (2017) use data from the US Social Security Administration to study the sources of lifetime inequality. They find that increases in lifetime inequality are primarily the result of increases in the variance of earnings at labor market entry rather than increases in the variance of life-cycle growth paths. Taking a structural approach, Huggett et al. (2011) reach a very similar conclusion using PSID data. My findings with German data are consistent with these results and suggest that increases in the variance of starting wages may result from changes in initial sorting. Additionally, growth in sorting at labor market entry suggests that the factors that generate individuals' initial productivity, such as education, childhood environment, or occupational choice, have a greater effect in determining the type of firm individuals are employed at over time. Also, the fact that reallocation is stronger at entry than over the life-cycle points to persistent sorting effects consistent with other studies that find

[8] See Eeckhout and Kircher (2011); Hagedorn et al. (2017); Lopes de Melo (2018); Bagger and Lentz (2018); Lise and Robin (2017).

persistent wage effects with respect to entry conditions (Kahn, 2010; Oreopoulos et al., 2012).

Next, I add to the literature studying the role of exports in sorting. Most theories of sorting are based on an assumption of complementarity in production between heterogeneous worker and firm types. In other words, high-skill workers are particularly productive at high-productivity firms. A classic result in a frictionless environment is that the optimal allocation is characterized by *assortative matching*, i.e. employment matches are symmetric such that the highest type worker matches with the highest type firm and the lowest type worker with the lowest type firm (Becker, 1973). In an environment with search frictions, each firm tolerates some deviation from the optimal allocation and, therefore, is willing to hire workers within a *matching set* (Shimer and Smith, 2000). Trade has been theorized to increase sorting by shrinking the size of the matching set so that the market allocation approaches the optimal allocation (Bombardini et al., 2015). Growth in export opportunities increases the size of the market, increases demand, and, hence, increases the value of output for any given match. As a result, firms have a higher willingness to pay to find an optimal match, shrinking the matching set and increasing sorting.[9]

A few recent studies have found empirical evidence for the connection between trade and sorting. Whereas as I focus on the effect of export exposure on sorting within German local labor markets, Davidson et al. (2014) find that reductions in export tariffs increase within-industry sorting in Sweden. They also find insignificant negative effects of increases in import tariffs. Their identification strategy relies on the fact that Sweden is a small country and, therefore, has limited ability to influence tariffs set by the European Union. The similarity of the Davidson et al. (2014) results, despite using data from a different country and employing a different identification strategy, suggest that the effect of exports on sorting may have applicability beyond West Germany.

[9] Davidson et al. (2008) and Helpman et al. (2010) construct trade models with labor market search that reach similar conclusions about the relationship between trade liberalization and sorting. The result that increased match output leads to a reduction in the matching set rests on the assumption of capacity constraints in hiring which lead to opporuntity costs in selection of the appropriate worker. See Eeckhout and Kircher (2011) for a detailed discussion.

Bombardini et al. (2015) study the effect of exports on sorting by assessing some predictions of their model using French data. They find evidence that exporting firms employ a more homogeneous workforce, which they interpret as evidence that exporting firms increase sorting in the labor market. I add a detailed analysis of the effects of trade on worker flows both within and between industries to understand how exporting leads to sorting. Consistent with trade theory, I find evidence that export exposure increases sorting within the manufacturing industry. However, I also find that export exposure has important effects on the service industry with the entry of low-wage workers to low-wage firms. My findings are also relevant to a literature studying the effects of trade exposure on long-term earnings and employment dynamics (Autor et al., 2014; Dauth et al., 2016; Müller et al., 2016), but which does not explicitly consider sorting.

Finally, I add to a small literature which seeks to quantify the sources of rising sorting. In their study of the effects of outsourcing on the labor market in West Germany, Goldschmidt and Schmieder (2017) estimate that outsourcing is responsible for approximately 8% of the rise in West German sorting.[10] Although this may be a lower bound due to a strict definition of outsourcing, the fact that trade from only Eastern Europe and China can account for 14% of the total rise in sorting suggests that exports are an important source of sorting. In addition, the worker flow channels through which trade affects sorting are similar to the aggregate channels. Therefore, more general increases in demand, including both domestic and international sources, may potentially account for a large share of the total rise in sorting.

[10]Goldschmidt and Schmieder (2017) compute counterfactual moments for the variance of establishment fixed effects and the covariance of establishment and worker fixed effects. Unfortunately, they do not state the value of the variance of worker fixed effects in their sample. Therefore, I take this value from Card et al. (2013) to compute the correlation of establishment and worker fixed effects noting that Goldschmidt and Schmieder (2017) perform the Abowd et al. (1999) estimation procedure with similar data and specification as Card et al. (2013).

1.2 Data

I use labor market data from the German Social Security system provided by the Institute for Employment Research (IAB) based on the Integrated Employment Biographies (IEB) datafile. Employers are required to submit a notice of employment for all employees and trainees subject to social security, which covers approximately 80% of all employment in West Germany. The major excluded groups are the self-employed and civil servants. Annual employment notifications include information on job duration and earnings for each employment spell. Hours of work are not disclosed, but employees are classified as part-time when working less than 30 hours per week.[11] The administrative data also includes the establishment identification number (EID), industry, and district of the establishment; and the individual identification number, year of birth, gender, education, and occupation of the individual. EIDs are uniquely assigned on the basis of ownership, municipality, and industry.

The main disadvantage of IEB earnings data is that it is censored at the highest level of earnings subject to social security contributions. This results in censoring of 10 to 14% of observations between 1985 and 2009 including about one third of white-collar workers (Card et al., 2013; Schank et al., 2007). Thus, I apply a Tobit wage imputation procedure following Card et al. (2013) and Dustmann et al. (2009) including lifetime and co-worker earnings variables.

Following Card et al. (2013), I restrict the sample to full-time employment in West Germany between the ages of 20 and 60 from 1985 to 2009. I limit earnings to one establishment per year and, therefore, select only the *main job*–defined as an employee's highest earning establishment. Employees in part-time and marginal jobs as well as trainees are excluded. I deflate all earnings to 2010 levels using the German CPI provided by Federal Reserve Economic Data (FRED).

I utilize a variety of datasets prepared by the IAB. For the analysis of earnings and employment

[11] Using the German Socio-economic Panel, Dustmann et al. (2009) provide evidence that the variance of hours worked was constant in West Germany after 1990.

I primarily use the Sample of Integrated Labour Market Biographies (SIAB) 1974-2010. The SIAB is a 2% random sample of individual employment histories and, thus, permits longitudinal analysis. To calculate a measure of sorting, I merge worker and establishment fixed effects estimated from the full IEB universe for West Germany in Card et al. (2013). Estimating the AKM methodology in small samples can finite sample bias. Hence I use fixed effects from Card et al. (2013) instead of estimating them with the 2% SIAB.

To construct trade shocks I use the United Nations Commodity Trade Statistics Database (Comtrade) which provides annual statistics of over 170 reporter countries detailed by commodities and partner countries. Using a correspondence between SITC rev 3 product codes and NACE 3-digit industry codes provide by Dauth et al. (2014), I translate trade flows from commodity to industry codes. In order to obtain an accurate count of employment shares within county-industry cells, I aggregate employment from a 50% sample of the IABs Establishment History Panel (BHP) which is an establishment level dataset covering the universe of German establishments subject to social security contributions.

For the purposes of this study, a LLM is defined as a *kreis* which roughly corresponds to a US county. The average population a West German kreis is approximately 200,000.

To illuminate firms' response to trade shocks, I use the Linked-Employer-Employee-Data from the IAB (LIAB) Longitudinal Model 1993-2010. The LIAB is based on a survey of establishments, known as the IAB Establishment Panel. This survey draws a stratified sample of establishments based on industry and establishment size, where large establishments are oversampled. Respondents are followed over time, creating a longitudinal account of annual sales and investment at the establishment level. Participation is voluntary, but the response rate is around 80% (Baumgarten, 2013). The LIAB is produced by merging IEB data with the IAB Establishment Panel by year and establishment.

1.3 Background on Sorting

1.3.1 AKM Model

I estimate agent types based on wage components from the Abowd, Kramarz and Margolis (1999) [AKM] regression model which captures fixed unobservable heterogeneity for both workers and firms through following regression equation:

$$y_{it} = \alpha_i + \psi_{j(i,t)} + x'_{it}\beta + r_{it} \tag{1.1}$$

for log wage y of individual i in year t. Worker heterogeneity is captured by the fixed effect α_i which represents the portion of an individual's earnings capacity that is fully portable across employers. Establishment heterogeneity is modeled with the fixed effect $\psi_{j(i,t)}$, where $j(i,t)$ is a function mapping workers to firms in each year. The establishment fixed effect captures a proportional pay premium common to all workers at establishment j. The fixed effects are identified off of wage variation induced by worker movements between firms. Time-varying worker characteristics x_{it} include year dummies and quadratic and cubic age terms interacted with educational attainment.[12] The controls for age by education effectively control for five separate experience gradients.[13] Although I suppress additional notation, this equation is estimated for four separate seven-year intervals p over the period 1985 to 2009. This flexibility allows worker and firm fixed effects to change over time within the same individual or firm.

Consistent estimation of the parameters of equation (1.1) relies on the standard OLS identification assumption of conditional orthogonality of the error term r_{it}. In this context, Card et al. (2013) argue that identification primarily rests on an assumption of *exogenous mobility*, i.e. conditional

[12]The linear term of age is omitted because it is not separately identified from the year effects.

[13]The Card et al. (2013) education groups are university, some college, apprentice, dropout, and missing.

on fixed effects and observable characteristics, movements between establishments are random. A controversial implication of this assumption is that job transitions are independent of worker-firm specific wage components that may arise, say, due to the presence of complementarities between worker and firm types in the production function.[14]

Although the exogenous mobility assumption is potentially restrictive, Card et al. (2013) provide evidence that this simple model fits the data well and is, therefore, a useful approximation of the wage equation. First, they estimate equation (1.1) with a interaction term η_{ij} in place of the worker and establishment effects. It is possible to identify either firm and worker effects or match effects, but not both at the same time. Although, the match effect model fits better, the reduction in the root mean squared error is relatively small–on the order of 10-15%–suggesting a limited role for match effects. Furthermore, they measure wage changes as workers move between firms with different average wages–a measure that does not rely on the AKM model structure. They show that wage gains from moving up the firm distribution are very similar to the wage losses from down the firm distribution. This symmetry result is consistent with the log separable form of the AKM estimation equation.[15]

In a different approach to the estimation of firm and worker heterogeneity in wage components, Bonhomme et al. (2016) simplify the space of potential firm effects into firm *classes*. They categorize each firm into one of ten classes based on its wage distribution using a clustering algorithm. They then treat firm classes as discrete fixed-effects and worker types as random-effects.

[14]Given that the wage equation is specified in terms of log wages, the model can be rationalized with a specific form of complemarties in which the production function of worker type x and firm type y is specified such that match output, $f(x, y)$, is equal to xy.

[15]Symmetric wage changes are also found in Portugal (Card et al., 2015) and the US (Song et al., 2018). However, there is also evidence against the exogenous mobility assumption. Abowd et al. (2017) propose a test of the exogenous mobility restriction. Using the Longitudinal Employer Household Dynamics (LEHD) data set, they reject the null hypothesis of exogenous mobility. Woodcock (2015) finds evidence for match effects using a mixed-effect estimator that allows correlation with time-varying observable worker characteristics, but requires the set of random effects to be orthogonal. Thus there is evidence of statistically significant match effects, but they do not appear to account for a large share of total wage variation.

This hybrid method reduces the number of worker and firm types and, therefore, enables the explicit estimation of interaction effects between firm and worker types.[16] Although their estimates suggest some departure from log additivity of firm and worker components, they find that log additive models, such as AKM, provide a very good approximation of the wage equation.[17] In other words, the interaction effects representing worker-firm specific components are not quantitatively significant. Therefore, with a variety of methods the AKM equation estimation has been found to approximate the wage equation well.

1.3.2 Identifying Sorting

I measure sorting as the correlation between AKM worker and establishment fixed effects. A distinction, however, must be made between measures of sorting based on wage components versus productivity components. Assortative matching on wage components has direct implications for inequality as high-wage workers work at high-wage firms. However, wage components may or may not represent productivity types. Other factors may affect firm wages besides productivity such as compensating differentials, bargaining strength, search frictions, or the opportunity costs of hiring workers given job scarcity. Therefore, sorting on wage components may reflect other factors besides worker-firm productivity complementarities.

The literature on sorting is largely focused on identifying the nature of complementarities in the production function in order to draw implications for aggregate efficiency. Earlier studies often interpreted negative or small correlations of AKM wage components as evidence against

[16]Card et al. (2013) estimate match effects, but not in the same regression as worker and firm effects. Estimating such a regression would require that every worker worked for every firm. By narrowing the type space, Bonhomme et al. (2016) can estimate a full model with separate effects for firm, workers, and their interactions.

[17]The R^2 increases from 74.8% in the model without interactions effects to only 75.8% with in the model with interaction effects (in the static version).

an interpretation of worker and firm types as complementary.[18] Subsequently, two key critiques were leveled against using AKM wage components to identify sorting of productivity types–one theoretical and the other empirical.

Eeckhout and Kircher (2011) caution against using AKM wage components to infer the sign of production complementarities due to theoretical inconsistencies between wage and production components. Using a simplified search model with positive assortative matching in the tradition of Shimer and Smith (2000) and Atakan (2006), they derive an analytical expression for the firm fixed effect. They show that the ranking of the firm fixed effect does not correspond to the ranking for firm productivity. Due to capacity constraints in the number of available jobs, firms face an opportunity cost of hiring a suboptimal worker type. If a worker is employed by a firm with lower productivity than her optimal match, she will earn less due to low match output. On the other hand, if a worker is employed by a firm with higher productivity than her optimal type, she will also earn less. In this case the firm must be compensated for the opportunity cost of not hiring its optimal worker type. As a result, workers experience a non-monotonic relationship between wages and firm productivity and achieve the highest wage at their optimal match.[19]

Although the results that follow can be strictly interpreted in terms of wages components, in settings where both employer-employee wage data and firm outcome variables are present, firm fixed effects are positively correlated with measures of firm productivity. Using German data, Card et al. (2013) find a positive correlation between firm fixed effects and firm survival. With Portuguese data, Card et al. (2015) find a significant and positive relationship between firm fixed effects and log value-added per worker. Using Swedish data, Davidson et al. (2014) find positive correlations of firm fixed effects and a variety of measures of firm productivity including labor

[18] See Abowd et al. (1999) for France; Abowd et al. (2002) for Washington State; Iranzo et al. (2008) for Italy; Gruetter and Lalive (2009) for Austria; Bagger and Lentz (2018) for Denmark; Lopes de Melo (2018) for Brazil; among others.

[19] In a related paper, Lopes de Melo (2018) reaches a similar conclusion.

productivity, size of capital stock, size of workforce, capital intensity, the ratio of R&D to sales, and the ratio of exports to sales.

The theoretical critique of AKM relies on a large role for opportunity costs in hiring as a result of capacity constraints. However, a model without capacity constraints can rationalize the AKM wage equation with sorting based on worker-firm production complementarities. To drive this point home, I construct a stylized model with these features in Appendix 1.11. I model the process of firm-worker matching through the firm's recruitment decision in an environment with search frictions and bargaining. Firms can increase their chance of finding a worker of a given type by increasing recruitment expenditure. I derive conditions for assortative matching based on the properties of the recruiting cost function. I show that complementarities between worker and firm types in recruitment costs must be stronger than complementarities in the production function to induce positive sorting. Thus high-type firms must face a lower cost of recruiting high-type workers. Such a feature can be rationalized through job referral networks or preferences over amenities that induce high-type workers to exert more search effort in finding high-type firms.[20] Therefore, the interaction of productive complementarities with other sorting mechanisms lead to assortative matching. The model features a wage equation that is log separable in firm and worker components. Furthermore, the wage components map to productivity types. See Appendix 1.11 for a full exposition of the model. [21]

In the end, the importance of these theoretical limitations depends on the magnitude of the opportunity cost of hiring a suboptimal worker type. If workers can be replaced easily or if firms have an unsatiated demand for labor, mismatch will not be costly, and, therefore the AKM ap-

[20]See Card et al. (2018) for an outline of a model which can produce sorting based on preferences for amenities by skill groups. See Schmutte (2014) for the role of job referral networks in facilitating the matching of high-ability workers to high-paying firms.

[21]Bagger and Lentz (2018) also present a model of sorting without capacity constraints. Due to the complexity of their approach, the wage equation is not strictly log separable. However, the model can be consistent with and AKM wage equation when workers have substantial bargaining power.

proach will provide an accurate approximation of the wage equation. To provide an estimate of the magnitude of opportunity costs, note that non-monotonic wage effects like those derived in Eeckhout and Kircher (2011) would load into worker-firm specific match effects. However, as previously noted, Card et al. (2013) show that match effects account for a modest share of wage variation and Bonhomme et al. (2016) find a limited role for work-firm specific components in a more flexibly specified model.[22]

The literature has also documented an empirical challenge in computing wage component sorting using the AKM components. Estimating the vast number of parameters in the AKM equation can produce a finite sample bias known as *limited mobility bias* (Andrews et al., 2008), which results when there are few job switchers per firm. Intuitively, as the two fixed effects roughly add up to total wages, deviations in the firm fixed effect caused by sampling error are counteracted by deviations in the opposite direction in the worker fixed effect. The result is a negative bias in the correlation of worker and firm effects caused by sampling error. In small samples or in samples with few movers, researchers often get negative estimates of the sorting. When researchers use larger samples, positive correlations are typically found (Card et al., 2013; Song et al., 2018). However, the level of sorting is not likely to be meaningful even in large samples. Therefore, I follow the literature and consider changes in sorting over time. This will work as long as the bias is stable over time. Card et al. (2013) provide evidence that the distribution of movers per firm is stable over time in the West German sample.

Due to a variety of estimates using both structural and more flexible econometric techniques, a consensus is emerging that there is positive sorting in the labor market. This is true in terms of sorting on both productivity and wage components.[23] Although there is strong evidence of positive

[22]Song et al. (2018) also reach a similar conclusion with US data.

[23]For structural estimates of positive sorting on productivity components see Hagedorn et al. (2017), Bagger and Lentz (2018), Lopes de Melo (2018), and Lise et al. (2016). A notable exception is Gulyas (2016). For estimates of positive sorting in wage components, see Card et al. (2013) and Song et al. (2018) using AKM on large datasets;

sorting, less attention has been given to how sorting is evolving over time. Relying on evidence that suggests the AKM approach is an accurate approximation of the wage equation, I measure changes in sorting based on the AKM wage components. In this approach, worker- and firm-types are estimated for each agent. The AKM approach, therefore, facilitates the analysis of the effects of worker flows on changes in labor market sorting.

1.3.3 Trends in Inequality and Sorting

West German wage inequality rose substantially over the three decade period from 1980 to 2010. Using German social security data, Dustmann et al. (2009) find a 0.6 log point annual increase in the 85/50 earnings ratio from 1975 to 2004. For perspective, the 90/50 ratio rose one log point per year over the same period in the US (Autor et al., 2008). Building on this result, Card et al. (2013) decompose the change in the variance of log wages into worker and firm components. Perhaps surprisingly, they find that firms contribute substantially to rising inequality.

Estimation of equation (1.1) allows for a straightforward decomposition of the variance of log wages into the following components:

$$\begin{aligned}\text{Var}\left(y_{it}\right) &= \text{Var}\left(\hat{\alpha}_i\right) + \text{Var}\left(\hat{\psi}_{j(i,t)}\right) + \text{Var}(x_{it}'\hat{\beta}) + \text{Var}(\hat{r}_{it}) \\ &+ 2\text{Cov}\left(\hat{\alpha}_i, \hat{\psi}_{j(i,t)}\right) + 2\text{Cov}\left(\hat{\alpha}_i, x_{it}'\hat{\beta}\right) + 2\text{Cov}\left(\hat{\psi}_{j(i,t)}, x_{it}'\hat{\beta}\right). \end{aligned} \tag{1.2}$$

Card et al. (2013) show that the primary drivers of the change in wage variance are the variance of worker effects (39%), the covariance of establishment and worker effect (34%), and the variance of establishment effects (25%). Therefore, the sorting of establishments and workers, as indicated by the covariance term, emerges as an important contributor to German wage inequality. In fact, in a

Bonhomme et al. (2016) and Borovicková and Shimer (2017) for econometric approaches that don't rely on AKM. Bartolucci et al. (2015) find positive sorting with worker wage components and firm profits.

counterfactural in which a more precise measure of sorting, the *correlation* between establishments and workers, is held constant they estimate that sorting accounts for 31% of the rise in total wage variance. Song et al. (2018) apply the AKM methodology to the US Social Security Administration data and find a strikingly similar result in that earnings variance rises to only 70% of its actual level between 1980 and 2013 with the correlation between firm and worker effects held constant.

1.3.4 Sorting in Local Labor Markets

I define sorting as the correlation of worker and firm fixed effects within a given local labor market (LLM) or:

$$Corr_l^p\left(\hat{\alpha}_i, \hat{\psi}_{j(i,t)}\right) \tag{1.3}$$

where l and p denote LLM and estimation interval, respectively. I compute a job-weighted correlation to represent sorting at the worker level.[24] Table 1.1 compares measures of national and LLM sorting. Columns (1) through (4) present the level of sorting in each interval while column (5) presents the change from the first to last interval. Despite a lower level of sorting, the changes in average LLM sorting are equal to the change in national sorting. The same cannot be said for average within-industry sorting. The rise in this measure is roughly half of the total rise. Women have a lower level and a smaller rise in sorting. However, within-LLM sorting is once again similar to the national change in sorting. Since the rise of LLM sorting mirrors the national rise, I use within-LLM measures to understand changes in national sorting.

[24]The statistic, therefore, describes how likely it is for a high-wage worker to work at a high-wage firm, as opposed to a firm-weighted measure that would describe the relationship between firms and their average worker fixed effect. Since the firm distribution is high-skewed such that there are a few very large firms and many very small firms, a firm-weighted measure places greater weight on small firms.

Table 1.1: Correlation of Establishment and Worker Fixed Effects over Time

		Int 1 '85-'91 (1)	Int 2 '90-'96 (2)	Int 3 '96-'02 (3)	Int 4 '03-'09 (4)	Change 1 to 4 (5)
	National	0.051	0.114	0.191	0.282	0.231
Male	Average Within-LLM	0.019	0.081	0.156	0.248	0.229
	Average Within-Ind	0.024	0.077	0.092	0.143	0.120
	National	0.042	0.088	0.097	0.137	0.095
Female	Average Within-LLM	0.022	0.063	0.071	0.111	0.088
	Average Within-Ind	-0.076	-0.030	-0.046	-0.012	0.064

Notes: "National" refers to the aggregate correlation of worker and establishment fixed effects over time the full sample. "Average Within-LLM" refers to a region-size weighted average of within-local labor market correlations of worker and establishment fixed effects. "Average Within-Ind" refers to the average within three-digit industry correlations of worker and establishment fixed effects.

1.4 Decomposition of Changes in Sorting into Worker Flows

In order to better understand the forces behind rising sorting in Germany, I decompose the total change in sorting into six categories based on worker transitions across seven-year intervals. The main forces of interest are *labor market entry*, *reallocation*, nonemployment, and *amplification*. For each flow, I compute the net effect defined as the difference between employment measures of entrants in the lead interval minus employment measures of exiters in the lag interval.

Labor market entry denotes a cohort effect measured as the difference between the employment status of labor market entrants relative to exiters. Entrants are defined as individuals who are below the minimum age threshold (20 years) in the first interval and then become employed in the second interval. Labor market exit is the reverse situation in which an employed worker in the first interval passes over the maximum age threshold (60 years) in the second interval. Therefore, labor market entry is in part reflects a mechanical relationship with age.

Reallocation refers to job-to-job transitions. This is a natural channel in which we may expect sorting to arise as it represents job switches for workers with high labor force attachment. I identify reallocation both within and between LLMs.

Nonemployment refers to net movements from of unemployment and a residual category called “other”. The administrative dataset includes accurate information on unemployment duration as the German SSA is also responsible for the administration of unemployment benefits. As long as a worker has previously been employed, the coverage of the unemployed population is almost universal. Entering flows from unemployment refer to individuals who are unemployed in the first interval, but become employed in the second interval. Exiting flows refer to individuals who transition from an employed to an unemployed state between intervals.

The employment category “other” refers to individual who are not unemployed but do not have a valid Card et al. (2013) fixed effect. This includes a variety of possible states including out of the labor force, employed in East Germany, part-time work, work in marginal jobs, self-employment, some civil servant employment, employed but not in the largest connected set of firms, immigration and emigration, and death. Analogously to the unemployment flows, I define entry as a transition from other to employment between the initial and subsequent intervals and exit as the inverse sequence. Given that the other category includes employment states, classification of this flow as an entry and exit flow versus a reallocation flow remains ambiguous.[25]

Job stayers denote the group of workers than stay employed at the same firm between intervals. This category quantifies the relative contribution of job stayers to changes in measured sorting. The correlation of establishment fixed effects (EFEs) and worker fixed effects (WFEs) can change due to changes in fixed effect values within a stable match. As the fixed effects are independently estimated in four intervals, I allow for this possibility. Amplification arises from changes in fixed effects that are correlated with initial conditions. For instance, suppose worker x is hired by firm

[25] Subsequent results provide evidence that this flow behaves more like a reallocation than a flow from unemployment.

y and both are at the top of their respective fixed effect distributions. If all worker-firm matches stay the same and all fixed effects are constant except that either the WFE of x or the EFE of y increases, then measured sorting will increase. Thus, amplification can arise due to either changes in WFEs that are correlated with initial EFE levels, changes in EFEs that are correlated with initial WFE levels, or changes in EFEs that are correlated with changes in WFEs.

1.4.1 Methods

The goal of the worker flow decomposition is to understand which worker flows are important for rising sorting. Sorting, defined as the average LLM correlation between WFE and EFEs, is a function of the joint distribution of WFE and EFEs. The idea, therefore, is to estimate an approximation of the joint distribution over time, and then perform counterfactual exercises in which worker flow channels are sequentially shut down. The worker flow components of sorting can be computed by taking the correlation over the counterfactual joint distributions. This method estimates the total effect of a given flow which may comprise both within and between worker-flow-group effects.

I approximate the joint density by computing WFE and EFE quintiles within LLMs in each period. The resulting joint distribution is a five by five grid of 25 cells designating employment shares. Define $\pi^p_{ij} \equiv \frac{E^p_{ij}}{E^p} \equiv \frac{\sum_{l=1}^{L} E^p_{ijl}}{\sum_{l=1}^{L} E^p_l}$ as the weighted average share of LLM employment in each WFE quintile i and EFE j and estimation interval p.[26] A measure of sorting for each period p across this simplified distribution can be computed as:

$$\rho^p \equiv \mathit{Corr}\left(\pi^p_{ij}\overline{\alpha}^p_i, \pi^p_{ij}\overline{\psi}^p_j\right) \tag{1.4}$$

for all i, j in $\{1, 5\}$, where $\overline{\alpha}^p_i$ denotes the average WFE in quintile i and $\overline{\psi}^p_j$ denotes the average

[26]Note that the weighted average share is simply the ratio of agreggate employments across LLMs as $\frac{1}{L}\sum_{l=1}^{L}\left[L\frac{E^p_l}{\sum_{l=1}^{L} E^p_l}\right]\frac{E^p_{ij}}{E^p_l} = \frac{\sum_{l=1}^{L} E^p_{ijl}}{\sum_{l=1}^{L} E^p_l}$.

EFE in quintile j. The change in sorting can be written as:

$$\triangle\rho \quad = \quad \mathit{Corr}\left(\pi_{ij}^{p+1}\overline{\alpha}_i^{p+1}, \pi_{ij}^{p+1}\overline{\psi}_j^{p+1}\right) - \mathit{Corr}\left(\pi_{ij}^{p}\overline{\alpha}_i^{p}, \pi_{ij}^{p}\overline{\psi}_j^{p}\right). \tag{1.5}$$

Note that the lead period employment share can be re-formulated as:

$$\begin{aligned} \pi_{ij}^{p+1} &= \left[\pi_{ij}^{p} + \frac{E_{ij}^{p+1} - E_{ij}^{p}}{E^p}\right]\frac{E^p}{E^{p+1}} \\ &= \left[\pi_{ij}^{p} + \frac{\triangle E_{ij}}{E^p}\right]\frac{E^p}{E^{p+1}}. \end{aligned} \tag{1.6}$$

In this expression, the initial period p employment share is added to the percentage change in employment for cell i, j. This sum is multiplied by a normalization term which accounts for total employment growth.

Let k denote worker flows between periods such that the sets E_{ijk} partition the sets E_{ij}. Then, $\triangle E_{ij} = \sum_k \triangle E_{ijk}$, where $\triangle E_{ijk}$ is the change in the total number employed in cell i, j from worker flow k. To compute counterfactual employment changes, I sequentially set employment changes to zero for each worker flow group. For example, the counterfactual employment for group $k = 1$ is computed as the change in employment shares when $\triangle E_{ij1} = 0, \forall\, i, j$ such that

$$\pi_{ij}^{p+1,C_1} = \left(\pi_{ij}^{p} + \frac{\sum_{k=2}^{6}\triangle E_{ijk}}{E^p}\right)\frac{E^p}{E^{p+1} - \sum_i\sum_j \triangle E_{ij1}}. \tag{1.7}$$

The final term of the expression re-normalizes the denominator to reflect counterfactual total employment in the second period. The counterfactual change in correlation is then defined as:

$$\triangle\rho^{C_k} \quad = \quad \mathit{Corr}\left(\pi_{ij}^{p+1,C_k}\overline{\alpha}_i^{p+1}, \pi_{ij}^{p+1,C_k}\overline{\psi}_j^{p+1}\right) - \mathit{Corr}\left(\pi_{ij}^{p}\overline{\alpha}_i^{p}, \pi_{ij}^{p}\overline{\psi}_j^{p}\right). \tag{1.8}$$

The contribution of each worker flow component is then simply computed as:

$$\triangle\rho - \triangle\rho^{c_k}. \tag{1.9}$$

This decomposition method is a partial equilibrium exercise since I assume that changes in the employment distribution do not affect the average fixed effect values $\overline{\alpha}_i$ and $\overline{\psi}_j$ or any of the other worker flows. Although restrictive, this is a standard assumption in decompositions of this type (DiNardo et al., 1995; Oaxaca, 1973; Blinder, 1973).

This is a generalization of the Oaxaca-Blinder approach. In my case the correlation is a function of six employment flow variables and ten prices (i.e. the average fixed effect values within the quintiles). To draw the analogy, let f be a general function of these arguments. Let e_k represent a vector of employment levels of all cells of the joint WFE-EFE distribution and e_{-k} all worker flows except k. In general terms, the contribution of flow k is equal to:

$$f\left(e_k^{p+1}, e_{-k}^{p+1}, \overline{\alpha}_i^{p+1}, \overline{\psi}_j^{p+1}\right) - f\left(e_k^{p}, e_{-k}^{p+1}, \overline{\alpha}_i^{p+1}, \overline{\psi}_j^{p+1}\right). \tag{1.10}$$

Essentially, I am taking an empirical derivative by holding all variables constant expect for the variable of interest. However, since the function f is not separable in all its arguments, this answer may change depending on the level of the variables being held constant. In practice, however, the levels of the other variables make very small differences to the final answer. In addition to decomposing the total change into changes in worker flows, I also account for changes in average quintiles prices. However, these components turn out to be small.

To implement this approach I apply a simple regression framework. Although these components can be computed as basic descriptive statistics without this framework, this setup creates continuity when I apply the decomposition method to estimates of the effect of trade on sorting. I estimate many weighted least squares regressions of the form:

$$\frac{\triangle E_{ijkl}}{E_l^p} = a_{ijk} + \delta^p + \epsilon_{ijkl}^p \tag{1.11}$$

where a_{ijk} is simply the parameter of a constant and δ^p is a fixed effect for period p. The regressions are weighted by initial labor market employment E_l^p. Given this weighting, $\hat{a}_{ijk}$ is the average change in worker -flow-cell employment as a proportion of total initial employment, or $\frac{\triangle E_{ijk}}{E^p}$. The aggregate estimate for the cell $\hat{a}_{ij}$ is similarly equal to the average change in cell employment divided by total initial employment, or $\frac{\triangle E_{ij}}{E^p}$. Thus, $\hat{a}_{ij} = \sum_k \hat{a}_{ijk}$.

Using this notation the expression for lead employment in equation (1.6) can be written as:

$$\pi_{ij}^{p+1} = \left[\pi_{ij}^p + \sum_k \hat{a}_{ijk}\right] \frac{E^p}{E^{p+1}}. \tag{1.12}$$

Counterfactual estimates involve shutting down each $\hat{a}_{ijk}$ sequentially and re-normalizing total second period employment, E^{p+1}, as in equation (1.7).

This method of approximating the correlation of worker and firm effects through quintiles works quite well. For example, the change in average LLM correlation between the estimation intervals 1985 to 1991 and 1996 to 2000 is 0.141 whereas the approximation method detailed above delivers a change of 0.139. For the second period, the approximation also works well with a 0.169 change in the average LLM correlation and an estimated 0.186 change from the quintile method. Therefore, it appears that little information is lost by approximating the distribution with quintiles.

1.4.2 Results

Table 1.2 presents the results of the decomposition exercise of changes in sorting into the contributions of worker flow groups. I report the results separately for each period. Columns (1) through

(4) present results for differences between the periods 1985 to 1991 and 1996 to 2002 and columns (5) through (8) for 1992 to 1996 and 2003 to 2009. For ease of exposition I refer to each period by its median year. For each period, the table is structured so that the first column, labeled E_k^p (%), presents the relative size of the worker flow group. The second column, labeled $\%\triangle E_k$, reports the growth in this flow relative to total initial employment. The third column, labeled $\triangle\rho_k$, reports the contribution of a given worker flow to sorting. The fourth column, labeled $\triangle\rho_k$ (%), reports the contribution of a given worker flow in relation to the total change in sorting.

Table 1.2: Contribution of Net Worker Flows to Changes in Corr(EFE,WFE)

	I. Interval 1: '88 to '99				*II. Interval 2: '93 to '06*			
	E_k^p (%) (1)	$\%\triangle E_k$ (2)	$\triangle\rho_k$ (3)	$\triangle\rho_k$ (%) (4)	E_k^p (%) (5)	$\%\triangle E_k$ (6)	$\triangle\rho_k$ (7)	$\triangle\rho_k$ (%) (8)
Labor market entry	30.4	-0.5	0.094	62.2	36.5	-9.1	0.108	51.8
Between-LLM job-to-job	*15.1*	*0.0*	*0.015*	*9.8*	*15.9*	*0.1*	*0.030*	*14.5*
Within-LLM job-to-job	*15.2*	*0.0*	*0.005*	*3.4*	*15.0*	*0.0*	*0.017*	*7.9*
Job-to-job	30.3	0.0	0.020	13.2	30.9	0.1	0.047	22.4
Other to emp.	*9.1*	*2.1*	*0.012*	*8.3*	*10.9*	*-0.6*	*0.020*	*9.6*
Unemp. to emp.	*5.4*	*-3.7*	*0.004*	*2.5*	*2.8*	*-1.0*	*0.005*	*2.6*
Nonemployment	14.6	-1.7	0.016	10.8	13.7	-1.6	0.025	12.2
Job Stayers	24.7	0.0	0.019	12.7	18.9	0.0	0.027	12.9
Change quintile values			0.002	1.1			0.002	0.7

Notes: Estimates of contributions of worker flows to LLM sorting based on the methology of Section 1.4.1. "E_k^p (%)" presents the initial share of a given worker flow relative to total LLM employment. "$\%\triangle E_k$" presents the change in employment of a given worker flow divided by initial total LLM employment. "$\triangle\rho_k$" presents the component of the change in the correlation of worker and establishment fixed that can be attributed to a given worker flow. "$\triangle\rho_k$ (%)" presents the contribution of a given worker flow as a share of the total change in sorting. "$\triangle$ avg. quintile vals" presents the contribution of changes in average quintile values across the establishment and worker fixed effect distribution to the total change in sorting.

The decomposition shows several interesting results. First, of the six worker flows, labor market entry is the most important determinant of increasing sorting across both periods. From 1988 to

1999, flows from net labor market entry throughout the fixed effect distribution were responsible for 62.2% of the total change. From 1993 to 2006, the corresponding number is 51.8%. These results represent the total effect of net labor market entry–potentially including a within- and a between-group component. The within-group component captures an effect in which labor market entrants are more sorted in their initial jobs than labor market exiters in their final jobs. The between group component is captured by differences in average WFE or EFEs between entrants and exiters.

The second most prominent channel is job stayers which comprises 12.7% and 12.9% of the total change in sorting in the first and second periods, respectively. This result suggests that changes in EFEs and WFEs are correlated with initial fixed effect levels. Therefore, contrary to the perception that sorting reflect movements of workers across firms and employment states, a measure of sorting based on the AKM methodology can increase when workers stay at the same firms.

Between-LLM job-to-job transitions also account for significant share of the total change in sorting–estimated at 9.8% and 14.5% for the two respective periods. Within-region reallocation appears to play a minor role–accounting for 3.4% and 7.9% in the two respective periods. Another important source of sorting is movements between other and employment. This channel accounts for 8.3% of the total effect from 1988 to 1999 and 9.6% from 1993 to 2006. These results suggest that over time workers hired out of "other" are more sorted that those exiting employment to "other". Although, this channel may identify individuals out of the labor force, these movements may also be job-to-job transitions from part-time, non-Social-Security-covered employment, or employment in East Germany. Adding the share from "other" to the shares from job-to-job transitions provides an upper bound for reallocation. The respective upper bounds in each period are 21.5% and 32.0%. The total effect of reallocation, therefore, although significant, is not the dominant force driving increases in sorting.

Net flows from unemployment to employment are responsible for a small share of the total change in sorting with 2.5% in the first period and 2.6% in the second. This result suggests that the firm that rehires an unemployed worker is similar to the firm that initially displaced her. Thus, this prominent labor market flow does not greatly affect sorting. This is partly due to the fact that unemployment transitions make up a small share of total worker flows.

Despite the fact that the AKM regression controls for returns to experience, there is a tendency for workers to move up the WFE distribution when they stay employed for consecutive periods. Therefore, if young workers enter with low WFE's in low EFE firms and subsequently move up both the WFE and EFE distribution over time, then the contribution of net labor market entry will include a life-cycle effect. In this case, comparing the net flows of exiters minus entrants may exaggerate the role labor market entry. For instance, suppose sorting remains unchanged between two periods, yet workers follow the life-cycle pattern described above. Then the total change in sorting is zero, but net labor market entry will report a positive contribution due to the life-cycle effect. Since the total effect sums to zero, other flows must reflect negative contributions. If this life-cycle effect is large, it can potentially complicate the interpretation of the aggregate results.

One way to remove the life-cycle component is to compare the contribution of net labor market entry across LLMs experiencing different rates of change in sorting. Assuming that the life-cycle effect is uncorrelated with the rate of change of sorting in a LLM will identify changes in sorting due only to differences in entry. This provides another reason to study the effect of trade on sorting. By comparing responses of different LLMs to trade shocks, the life-cycle effect is differenced out.

1.5 The Impact of Trade Shocks on Labor Market Sorting

As documented in Dauth et al. (2014), Germany experienced a surge in trade flows to and from both Eastern Europe and China from 1990 to 2010. The growth in trade corresponds to the opening of China and the fall of the Soviet Union in Eastern Europe. Both of these events can be viewed as largely exogenous to domestic German industry and, therefore, serve as useful shocks to analyze the effects of trade. In additional to their initial openings, both of these regions joined the WTO around 2001 which led to a further economic integration.

1.5.1 Methods

The work of Autor et al. (2013) has become influential in the study of trade on local labor markets (LLMs) and I largely follow their methodology. I construct a Bartik-style measure of regional export exposure per worker by assigning national changes in industry exports to local labor markets based on their initial share of industry employment:

$$\triangle EXP_{lt}^{GER} = \sum_{s} \frac{E_{lst}}{E_{st}} \frac{\triangle EXP_{st}^{GER \rightarrow EAST}}{E_{lt}}. \tag{1.13}$$

$\triangle EXP_{st}^{EAST}$ denotes the observed change in national exports from Germany to the East between time period t and $t+1$ in industry s. E_{lt} is total employment in region l in period t. E_{lst}/E_{st} is the share of national employment of industry s employed in region l in initial period t. I create a measure of import exposure $\triangle IMP_{st}$ with the corresponding equation using national imports.

Although the opening of the East to trade can be viewed as exogenous to domestic industry at the moment of initiation, the gradual and continuous nature of this process warrants the use of an additional instrument to disentangle potentially endogenous supply and demand factors emerging over time. In other words, over short periods of time around the initial opening to trade and accession to the WTO, we may expect trade flows from the East to be exogenous across industries in the domestic German market. However, over time industries may evolve such that trade flows represent endogenous industry supply differences. For example, suppose that the German car industry innovates successfully to capture global market share. Differences in industry trade flows to the East will in part reflect these innovations of the car industry, rather than the pure demand effects of market access.

To alleviate concerns over endogeneity, I follow Autor et al. (2013) and Dauth et al. (2014) by constructing an instrument based on trade flows to nine comparably developed countries which are not members of the European Monetary Union:

$$\triangle \mathit{EXP}^{Other}_{lt} = \sum_{s} \frac{E_{ls,t-1}}{E_{s,t-1}} \frac{\triangle \mathit{EXP}^{Other \rightarrow EAST}_{st}}{E_{l,t-1}}. \tag{1.14}$$

In addition the trade flows to other countries, the instrument varies from the measure of trade exposure as employment levels are measured with ten year lags. The countries upon which the instrument is constructed are Australia, Canada, Japan, Norway, New Zealand, Sweden, Singapore, and the United Kingdom. In the case of exports shocks, the instrument is constructed to isolate the effect of foreign demand from domestic supply. Using variation based on goods that many other developed countries buy from and sell to the East, removes idiosyncratic Germany supply and demand components. The instrument is based on lagged industry employment shares to alleviate the concern that some industries could anticipate high returns from Eastern market access and, therefore, mobilized ahead of time.

I follow Dauth et al. (2014) and Autor et al. (2013) in analyzing the effect of export and import shocks on local labor market outcomes by estimating the following equation:

$$\triangle y_{lt} = \beta_1 \triangle \mathit{EXP}_{lt} + \beta_2 \triangle \mathit{IMP}_{lt} + \gamma X_{lt} + \lambda_{r(l)} + \delta_t + \epsilon_{lt} \tag{1.15}$$

where y_{it} represents a labor market outcome of local labor market l in time period t, X_{lt} represents initial labor market characteristics of a region, $\lambda_{r(l)}$ represents a region fixed effect in which $r(l)$ denotes a function from counties to larger geographic regions, and δ_t captures a time period fixed effect. Since the regression is performed in changes, the fixed effects capture common trends rather than levels. Changes are denoted as $\triangle$, such that $\triangle y_{lt} = y_{lt+1} - y_{lt}$.

1.5.2 Results on Employment and Wages

Prior to presenting the main results of the effect of trade on sorting, I present evidence that trade indeed acts as a demand shock to the affected industries. Conceptually the idea is that trade works through demand to induce sorting. Although I have no direct measure of demand, I present evidence on the effect of trade on manufacturing employment and wages.

Table 1.3 presents the results of estimating equation (1.15) with four separate dependent variables. The first column estimates the effect of trade shocks on manufacturing employment. Export exposure significantly increases employment with a coefficient of 1.358 log points. At the mean level of export exposure, this represents a 5.4% increase in manufacturing employment. Import exposure results in a similar decline of employment. Wages, on the other hand, are a different story. Export exposure is estimated to increases wages by 0.334 log points while there is an insignificant decline in wages from import exposure. At the mean level of export exposure, this represents 1.3% increase in manufacturing wages. An increase in employment coupled with rising wages is indicative a shift upward in demand for manufacturing labor. Thus, I interpret export shocks as increasing labor demand via product demand.

Columns (3) and (4) represent the establishment fixed effect (EFE) and worker fixed effect (WFE) components of wages, respectively. While neither, component is significant for imports, export exposure leads to an increase in the average level of WFEs for workers employed in manufacturing firms. Therefore, on average, the increase in wages is realized through an change in composition rather than an increase in the firm wage premium.

1.5.3 Results on Sorting

My primary outcome of interest is the correlation of EFEs and WFEs within a LLM. As fixed effects represent a component of wages, an increase in the correlation of the two components represents an increase in inequality as high wage establishments are more likely to employee high

Table 1.3: 2SLS Results of Employment and Wages on Trade Shocks

	△ Emp (1)	△ Wage (2)	△ EFE (3)	△ WFE (4)
Export exposure	1.358***	0.334**	-0.072	0.435***
	(0.430)	(0.160)	(0.157)	(0.116)
Import exposure	-1.519***	-0.098	0.000	-0.042
	(0.585)	(0.230)	(0.126)	(0.180)
Labor market controls	Y	Y	Y	Y
# geo f.e.'s	214	214	214	214
Adj R^2	0.711	0.599	0.995	0.993
N (county-periods)	650	650	650	650

Notes: All 2SLS regressions are weighted by the initial size of the regional labor force. Standard errors are clustered at the LMR2 level. Labor market controls include: initial level of sorting, % employment in manufacturing, % high skilled, % foreign born, % female, and % routine occupation.

wage workers.

Table 1.4 presents estimates of the causal effect of trade shocks on male labor market sorting. I estimate equation (1.15) with multiple specifications, primarily varying the geographic fixed effect. Column (1) presents the OLS estimates which indicate that export exposure shocks intensify sorting whereas import shocks have the opposite effect. The magnitude of the effect of export exposure is larger and is estimated with smaller standard errors. All regressions are weighted by initial county employment so that the results represent average worker-weighted effects across LLMs. Column (2) presents IV estimates without controls. Column (3) adds controls for state trends in sorting. Following Autor et al. (2013) and Dauth et al. (2014), I also include controls for the initial characteristics of the labor market including: percentage of employment in manufacturing, percentage of high skilled employment, percentage of foreign born employment, percentage of female employment, and percentage of routine occupation employment. These controls reduce the magnitude of both coefficients and remove any statistical significance from the import coefficient.

Table 1.4: 2SLS Results of Sorting on Trade Shocks for Males

	Region fixed effect					
	OLS: None (1)	IV: None (2)	IV: State (3)	IV: LMR1 (4)	IV: LMR2 (5)	IV: LMR2 (6)
Export exposure	0.0093***	0.0131***	0.0109***	0.0120***	0.0118***	0.0080***
	(0.0022)	(0.0029)	(0.0033)	(0.0028)	(0.0029)	(0.0025)
Import exposure	-0.0028	-0.0082*	-0.0047	-0.0045	-0.0020	-0.0014
	(0.0025)	(0.0042)	(0.0043)	(0.0046)	(0.0055)	(0.0050)
Initial sorting						-0.6185***
						(0.0050)
Labor market controls	N	N	Y	Y	Y	Y
# geo fixed effects	0	0	11	74	214	214
Adj R^2	0.093	0.076	0.115	0.212	0.278	0.445
N (county-periods)	650	650	650	650	650	650

Notes: All 2SLS regressions are weighted by the initial size of the regional labor force. Standard errors are clustered at the LMR2 level. Labor market controls include: % employment in manufacturing, % high skilled, % foreign born, % female, and % routine occupation.

The export coefficient, however, remains highly significant.

Columns (4), (5), and (6) use a taxonomy of labor market regions produced by Dustmann and Glitz (2015) denoted as *LMR1* and *LMR2*, respectively. The variation in labor market outcomes is at the county level of which there 325 in West Germany. LMR1 and LMR2 represent broader definitions of labor markets which resemble commuting zones in the US. There are 74 LMR1s and 214 LMR2s in West Germany. A potential concern is that the estimates reflect long-term industry trends rather than the changes in terms of trade with the East. If the geographic concentration of industries is stable, then using within-region variation over time alleviates these concerns. As the geographic fixed effects become narrower, the estimates use less cross-sectional variation and produce a more conservative test.

The export coefficient stays stable as successively less variation is used. The import coefficient, however, declines significantly and loses statistical significance. Column (6) add a control for the initial level of labor market sorting. This variable is significant and reduces the export coefficient slightly. As the most conservative specification, Column (6) represents the main specification for the analysis. Appendix Table 1.10 shows that similar results hold for women.[27]

The economic magnitude of the export coefficient is significant. Multiplying the export and import coefficients of Column (6) by the average change in trade exposure over the period 1988 to 2008, the total predicted change in the correlation is 0.0392.[28] The total change in average within-LLM sorting over the sample period is 0.229.[29] The predicted change represents 17.1% of the total change in sorting. However, following Autor et al. (2013) I use a more conservative estimate the total change in trade. Given that the estimated effect is with respect to the exogenous change in

[27]Add robustness results for separate intervals and differences between Eastern Europe and China.

[28]The employment weighted average change in exports (imports) is 7.61 (6.25) from 1988 to 2008. Therefore, 7.62*0.0080-6.25*0.0014=0.0522. Given that the intervals overlap I think multiply by 0.75 to get 0.0392

[29]See Table 1.1.

trade, I use only the proportion the variation in trade exposure explained by the instrument. This allows for the possibility that endogenous changes in trade flows do not have an effect on sorting. Therefore, I estimate that trade with the East is responsible for 0.0326 or 14.2% of the total change in West Germany sorting.[30]

1.6 Decomposition of the Effect of Trade on Sorting into Worker Flows

1.6.1 Methods

To decompose the effect of trade on sorting into worker flow channels I follow a similar method as described in Section 1.4.1. However, instead of running a simple regression to compute the average change in employment cells as in equation (1.11), I estimate the full trade model of equation (1.15). I estimate the effect of trade shocks on employment changes throughout the joint WFE-EFE distribution:

$$\frac{\triangle E_{ijkl}}{E_l^p} = \beta_1^{ijk} \triangle EXP_{lt} + \beta_2^{ijk} \triangle IMP_{lt} + \gamma^{ijk} X_{lt} + \lambda_{r(l)}^{ijk} + \delta_t^{ijk} + \epsilon_{lt}^{ijk}. \tag{1.16}$$

The dependent variable in equation (1.16) is the change in employment in WFE-EFE cell ij, worker flow k, and LLM l divided by total initial employment in LLM l. It therefore represents the change in a given employment cell as a share of total initial LLM employment. This equation is estimated by two-stage least squares using the same instrument as described in Section 1.5.1.

[30] For exports the ratio of instrument to total variation is 0.83, for imports 0.82. Therefore, the calculation becomes 0.75(7.62*0.0080*0.83-6.25*0.0014*0.82)=0.0326.

Similar to equation (1.12) I compute the total change in employment due to export shocks as:

$$\pi_{ij}^{p+1} = \left[\pi_{ij}^{p} + \sum_{k} \hat{\beta}_1^{ijk}\right] \frac{E^p}{E^p + \hat{\beta}_1} \tag{1.17}$$

where $\hat{\beta}_1 = \sum_k \sum_i \sum_j \hat{\beta}_1^{ijk}$. Counterfactual distributions are computed by shutting down the worker flow channels of trade, $\hat{\beta}_1^{ijk}$, sequentially and adjusting the denominator of the population adjustment factor appropriately.

The approximation to quintiles works well in this case also. The decomposition method yields an total change in sorting due to exports of 0.0093. This is similar to the regression coefficient of 0.0080.

1.6.2 Worker Flows

Table 1.5 presents the results of a decomposition of the effect of export exposure on sorting into worker flow channels. Panel I presents the results of the worker flow decomposition. Column (1), labeled $\triangle\rho_k$, presents the contribution of a given flow to changes in sorting. Column (2), labeled $\triangle\rho_k$ (%), presents the worker flow contribution as a percentage of the total change. Panel II presents a picture of the size of each worker flow channel. Column (3), labeled E_k^p (%), presents the initial share of a given worker flow as a percentage of total initial employment. Column (4), labeled $\%\triangle E_k$, presents an estimate of the export-induced change in employment of a given worker flow as a percentage of total initial employment. Panel III takes an average of the worker flow share contributions from Table 1.2.

Turing to the results in Panel I, we see that once again the most important contributor to increased sorting is flows into and out the labor market. Net labor market entry alone comprises 47.7% of the total effect of exports on sorting. This net effect is smaller than the contribution to the aggregate effect of sorting. This is consistent with the idea that some portion of the net labor market entry share of the aggregate sorting change is due to a life-cycle component. However, this component is not large enough to change the general conclusion that worker flows into employment from labor market entrants are the most significant contributors to changes in LLM sorting.

Table 1.5: Decomposition of Export Sorting into Worker Flows

	I. Components of Change in Sorting through Exports		*II. Employment Shares*		*III. Components of Change in Aggregate Sorting*
	$\triangle\rho_k$ (1)	$\triangle\rho_k$ (%) (2)	E_k^p (%) (3)	$\%\triangle E_k$ (4)	$\triangle\rho_k$ (%) (5)
Labor market entry	0.0045	47.7	33.5	0.35	57.0
Between-LLM job-to-job	*0.0016*	*16.6*	*15.5*	*0.18*	*12.1*
Within-LLM job-to-job	*0.0000*	*0.0*	*15.1*	*0.00*	*5.7*
Job-to-job	0.0016	16.6	30.6	0.18	17.8
Other to emp.	*0.0009*	*9.8*	*10.0*	*0.20*	*8.9*
Unemp. to emp.	*0.0000*	*0.3*	*4.1*	*0.17*	*2.6*
Nonemployment	0.0009	10.1	14.1	0.37	11.5
Job Stayers	0.0023	25.0	21.8	0.00	12.8

Notes: Estimates of contributions of worker flows to export-induced LLM sorting based on the methology of Section 1.6.1. "$\triangle\rho_k$" presents the component of the change in the correlation of worker and establishment fixed that can be attributed to a given worker flow through export exposure. "$\triangle\rho_k$ (%)" presents the contribution of a given worker flow as a share of the total export-induced change in sorting. "E_k^p (%)" presents the initial share of a given worker flow relative to total LLM employment. "$\%\triangle E_k$" presents estiamtes of the *export-induced* change in employment of a given worker flow divided by initial total LLM employment.

Job stayers comprise the second most important flow–contributing 25.0% of the total export effect. This result suggests that demand shocks, at least when targeted toward manufacturing firms, produce increases in wage components that are correlated with the initial WFE-EFE distribution.

Reallocation is also a significant contributor to sorting as job-to-job transitions account for 16.6% of the total effect. Interestingly, within-LLM reallocations account for none of the total effect as it is all driven by between-LLM job switching. Net flows into and out of the other category are also a significant contributor to sorting–accounting for 9.8% of the total export effect. The total contribution of reallocation can be bounded at 26.4%. Thus, at most, about a quarter of the increase in sorting following export shocks are driven by job-to-job transitions.

A striking result emerges through a comparison of the trade-induced sorting flows with the aggregate sorting flows. Column (5) presents the average contribution of each worker flow as a percentage of the total change in sorting across both periods. The aggregate shares are quite similar to the trade shares. The main differences are that net labor market entry is slightly less important in the trade flows and both between-LLM reallocation and job stayers are slightly more important. These results suggest that increases in demand from other sources, whether domestic or international, have the potential to explain a larger share of the total change in sorting than the direct effect of trade from the East alone.

1.7 Export-Induced Worker Flows by Industry and Firm Type

1.7.1 Connection of Results to Trade Theory

Despite its simple form, it is difficult to incorporate all the features of the AKM empirical structure into a theoretical model. The existence of firm fixed effects requires a departure from perfect competition in the labor market to allow for the possibility that similar workers are paid differently depending on their place of work. Variance in firm premiums requires a source of firm heterogeneity. A common approach used to incorporate firm heterogeneity is to allow heterogeneity in productivity and monopolistic competition in the product market. Finally, variance in worker fixed effects requires heterogeneity in worker productivity. All told a theoretical model would require productive heterogeneity on both the firm and worker side as well as imperfect market competition in both the product and labor market. To the best of my knowledge, no current trade model is able to incorporate all of these features.

Trade models of within-industry firm heterogeneity typically abstract from modeling the interactions of geographic- and industry-specific labor markets. Given that I measure the full effects of local labor market (LLM) and manufacturing sector specific export shocks, I allow for the possibility of between-LLM and between-industry reallocation.[31] These channels provide alternative margins of adjustment which are important to consider when applying the predictions of the trade literature to the results.

In lieu of a fully specified trade model that is consistent with the AKM structure, I briefly describe some important predictions of the recent trade literature to guide the interpretation of the subsequent results. The model of Melitz (2003) has become influential in the study of international trade through its treatment of firm heterogeneity in productivity. A key prediction of this model is that trade liberalization disproportionately benefits the most productive firms as increases in demand accentuate productivity differences. As top firms bid up the price of labor, wages rise for all firms. As a result the least productive firms exit. In the end, market share is reallocated from the least to most productive firms–raising aggregate efficiency. Note that despite the fact that the trade shock is common across the industry, heterogeneous responses impact within-industry variation in productivity.

Sampson (2014) extends the Melitz (2003) model by allowing for worker heterogeneity in skill and endogenous technological choice on the firm side. By incorporating heterogeneity on both sides of the labor market, Sampson (2014) connects the literature on assortative matching to a trade context. An assumption of complementaries in the production function leads to positive assortative matching–a standard result. Trade liberalization has similar implications as in the Melitz (2003) model, as the most productive firms profit most. However, is this case dispersion in firm

[31] Autor et al. (2013) find strong wage effects in the non-manufacturing sector as a result of the China import shock in US LLMs.

productivity leads to dispersion in worker wages. Since top workers work at top firms, marginal product of labor increases more for high type workers and hence wage inequality increases. If firms can choose their optimal level of technological investment, trade liberalization induces firms to upgrade technology. Given a closed labor market, however, the matching of firms and workers is constant. Extending the model to allow firms to hire workers from other industries or regions could change this result. Both demand shocks and technological upgrading would provide an incentive for firms to upgrade the quality of their workforce.

Some predictions broadly consistent with this strand of the trade literature are as follows.[32] The most productive manufacturing firms benefit the most from export exposure. Given the presence of between-industry and between-region margins of adjustment, top manufacturing firms also upgrade the quality of their workforce. Furthermore, manufacturing firms increase technological investment as a result of trade liberalization to fully exploit their revenue potential. Insofar as technology is complementary with high-skill labor and a substitute for low-skill labor, we expect a relative reduction of low-skill labor in manufacturing firms as a result of demand increases.

1.7.2 Export Flows by Industry

In order to clarify the effects of the export shock I classify firms into broad industry sectors: manufacturing and non-manufacturing. Table 1.6 presents the results of the decomposition of the effect of trade on sorting into work flows by sector. Panel I presents the results of a decomposition while panel II shows the initial employment shares and the estimated total change in employment for a given worker flow group.

A striking result is that all of the effect of net labor market entry works through the non-

[32] I am considering the case in which the rise of the East acts as either an inrease in trading partners or decrease in the variable costs of trade. Reductions in the fixed costs of trade will have subtely different predictions. See Melitz (2003) and Sampson (2014) for a discussion of the comparative statics.

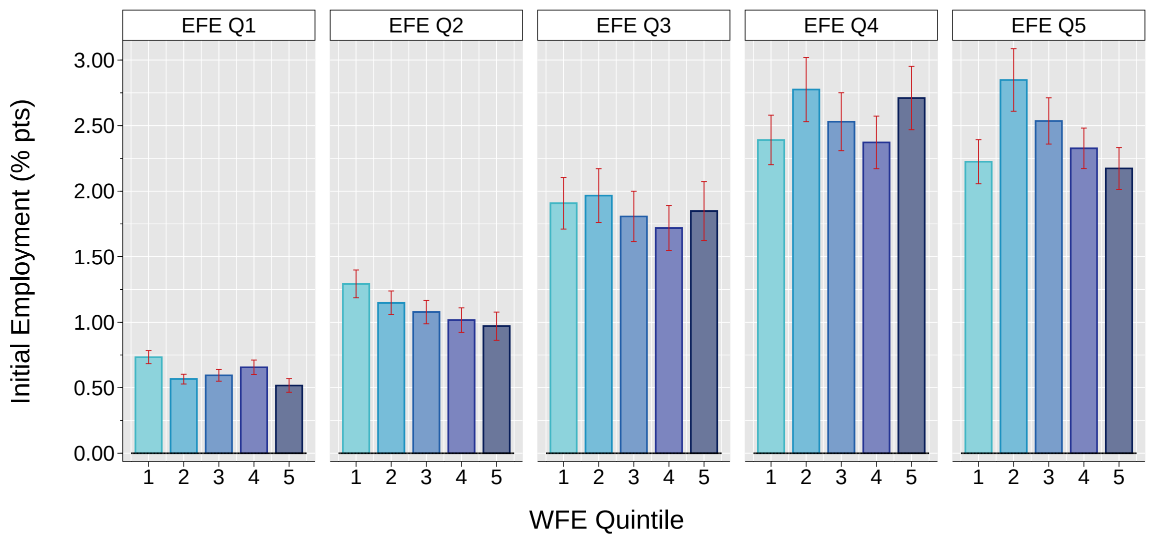
Table 1.6: Decomposition of Export Sorting into Worker Flows by Industry

	I. Components of Change in Sorting through Exports				*II. Employment Shares*			
	Manufacturing		Non-Manufacturing		Manufacturing		Non-Manufacturing	
	$\triangle\rho_k$ (1)	$\triangle\rho_k$ (%) (2)	$\triangle\rho_k$ (3)	$\triangle\rho_k$ (%) (4)	E_k^p (%) (5)	$\%\triangle E_k$ (6)	E_k^p (%) (7)	$\%\triangle E_k$ (8)
Labor market entry	0.0002	2.4	0.0042	45.7	13.7	-0.14	19.8	0.49***
Between-LLM job-to-job	*0.0008*	*8.4*	*0.0008*	*8.3*	*5.5*	*0.18****	*9.9*	*0.00*
Within-LLM job-to-job	*0.0003*	*3.1*	*-0.0003*	*-3.1*	*6.6*	*0.00*	*8.5*	*0.00*
Job-to-job	0.0011	11.5	0.0005	5.2	12.1	0.18	18.5	0.00
Other to emp.	*0.0007*	*8.0*	*0.0002*	*1.9*	*3.4*	*0.13*	*6.5*	*0.07*
Unemp. to emp.	*-0.0002*	*-2.1*	*0.0002*	*2.4*	*1.9*	*0.20****	*2.3*	*-0.04***
Nonemployment to emp.	0.0005	5.8	0.0004	4.3	5.3	0.33	8.8	0.04
Job Stayers	0.0029	30.7	-0.0005	-5.6	11.6	0.00	10.2	0.00
Industry total	0.0047	50.5	0.0046	49.5	42.7	0.37	57.3	0.52

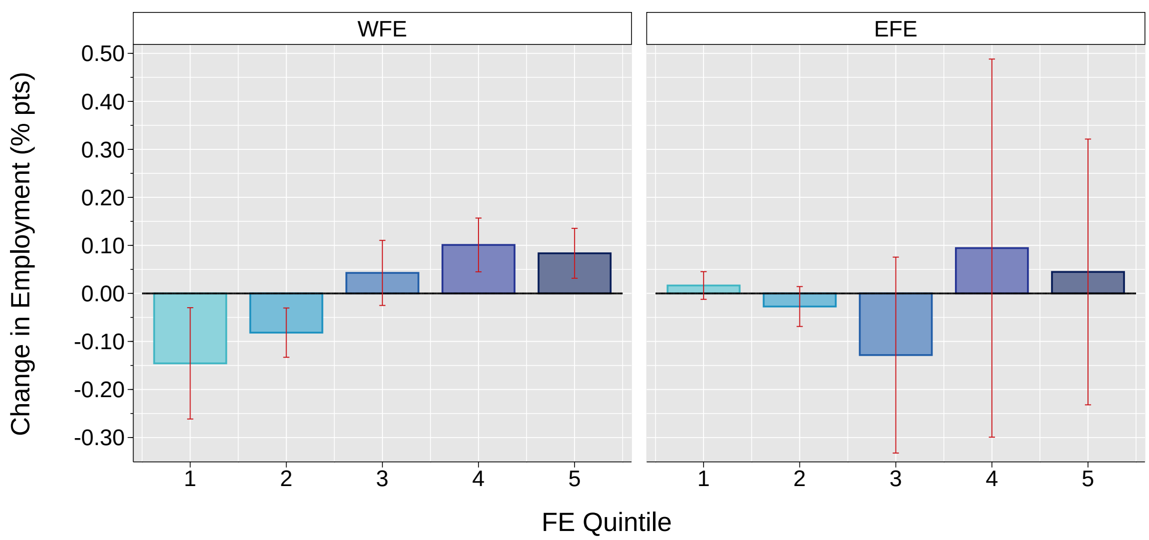
Notes: Estimates of contributions of worker flows to export-induced LLM sorting based on the methology of Section 1.6.1. "$\triangle\rho_k$" presents the component of the change in the correlation of worker and establishment fixed that can be attributed to a given worker flow through export exposure. "$\triangle\rho_k$ (%)" presents the contribution of a given worker flow as a share of the total export-induced change in sorting. "E_k^p (%)" presents the initial share of a given worker flow relative to total LLM employment. "$\%\triangle E_k$" presents estiamtes of the *export-induced* change in employment of a given worker flow divided by initial total LLM employment.

Figure 1.1: Amplification in the Manufacturing Sector

(a) Initial Joint WFE-EFE Distribution of Manufacturing Jobs



(b) Export-Induced Changes in FEs for Job Stayers in Manufacturing Firms



Notes: Red brackets represent 90% confidence intervals. Subfigure (a) plot the intial distribution of manufacturing employment across the joint WFE-EFE distribution by taking an average of the intial distributions of both periods of change: ‘88 to ‘99 and ‘93 to ‘06. Subfigure (b) presents the estimated coefficients from equation (1.15) using total employment changes in each quintile of the marginal EFE and WFE distributions divided by initial total LLM employment.

manufacturing sector. The first row of column (2) shows that changes in net labor market entry patterns into manufacturing firms explain only 2.4% of the total export effect. On the other hand, changes in net labor market entry into the non-manufacturing sector are responsible for 45.7% of the total change in export-induced sorting. This is despite the fact that these flows comprise only 20% of the initial employment distribution (column (7)). Furthermore, comparing columns (6) and (8) we can see a difference in the total flows accruing to each sector. Entrants to export exposed LLMs are significantly *less* likely to enter the manufacturing sector and *more* likely to enter the non-manufacturing sector. The net effect is insignificantly different from zero. Therefore, entrants are not in general more likely to enter the labor market in export exposed LLMs, but there is a significant switch away from the manufacturing sector despite the positive demand shock.

Another interesting result is a stark divergence in the effect of job stayers on sorting in the manufacturing versus the non-manufacturing sector. The effect of job stayers is fully accounted for through the manufacturing sector. In fact the effect of job stayers on sorting in non-manufacturing firms is slightly negative. In so far as amplification is the result of increases in within-type fixed effects, the result that amplification works mainly through the manufacturing sector is consistent with expectations. After all, the direct effect of the export-induced demand shock is to the manufacturing sector. Figure 1.1 presents a visual representation of amplification in the manufacturing sector. Subfigure 1.1a shows the initial distribution of manufacturing jobs. Each panel represents the share of total LLM employment in each EFE quintile across the range of WFE quintiles. Manufacturing jobs are disproportionately represented in the top two quintiles of the EFE distribution. Subfigure 1.1b reports estimated coefficients of the effect increases in export exposure on change in employment across the marginal distributions of WFE and EFEs. Export-induced demand shocks tend to increase WFEs but not EFEs. Therefore, these results combine to offer a picture of amplification in which jobs with initially high EFEs respond to demand shocks in export-exposed LLMs through increases in WFEs.

The result that demand shockss lead to increases in WFEs of jobs stayers suggests manufacturing workers receive a wage increase that is portable across industries. This result is consistent with a competitive labor market in which the skill of manufacturing workers are substitutable across industries. This result may stem from the fact that the LLM demand shock is derived from a national industry shock. We may expect the firm specific component of wages, the EFE, to play a larger role with firm- or region-specific shocks.

In contrast to the story for labor market entry and amplification, the industry effects of between-region reallocation are roughly equal. Column (2) shows that between-LLM job-to-job transitions account for 8.4% of the total effect. Column (4) shows that regional reallocations in the non-manufacturing sector account for 8.3%. Adding net "other" and net within-LLM flows produces an upper bound of 19.5% for manufacturing reallocation and 7.1% for non-manufacturing reallocation.

Columns (2) and (4) of the final row report the total contribution of each industry to export-induced sorting. At 50.5% for manufacturing versus 49.5% for non-manufacturing, the share are roughly equal. Given that the decomposition method aggregates between- and within-industry effects, it remains unclear which if these effects arise due to within-industry sorting. As previously noted, the literature emphasizes how export exposure can lead to within-industry dispersion in productivity which can lead to within-industry wage dispersion. To test whether this industry components represent within-industry sorting effects, I estimate the trade exposure equation (1.15) but instead use within-industry/within-LLM correlation of fixed effects as the dependent variable. Using the main specification, I estimate a coefficient on the effect of export exposure on within-manufacturing industry sorting of 0.0083 with a standard error of 0.0050 which is statistically significant at the 10% level. Although not as precisely estimated as the aggregate coefficient, this result provides evidence consistent with a view that industry demand shocks cause within-industry sorting. Furthermore, the magnitude of the coefficient is very similar to the aggregate coefficient.

For the non-manufacturing industry, I estimate a positive, but statistically insignificant coefficient of export exposure on within-non-manufacturing sorting of 0.0041. This suggests that the effects of the non-manufacturing industry on export sorting are mostly between-industry effects.

1.7.3 Export Flows by Firm Type

Guided by trade theory, I analyze the results of export shocks on the sorting components across the EFE and initial firm size distributions. I use both EFEs and firm size as a proxy from firm productivity. Firm size is a measure which is correlated with productivity in many search and trade models. Section 1.3 provides evidence that EFEs are positively correlated with measures of firm productivity in a variety of settings.

The simplicity of my decomposition method allows it to flexibly cover a wide range of potential counterfactuals. In theory, I can estimate the effect of changes in each cell of the WFE, EFE, worker flow, and industry distribution on the total change in sorting. In the subsequent analysis, I decompose the contribution of worker flows at different points in the EFE distribution. Specifically, I compute separate counterfactual shares for three firm groups: low-, mid-, and high-EFEs.[33] This exercise provides a picture of the location in the joint distribution by which each flow affects sorting. For instance, a large share for the low-EFE group means that a change in employment for low-EFE firms increased sorting. In order for low-EFE firms to increase sorting it must be the case that they gained relatively more low-WFE workers.

In order to provide a more complete picture of within-industry reactions to export shocks, I also condition by the size of the firm in the initial period. Within each LLM and sector, I compute worker-weighted firm size medians. I categorize firms as either small continuing, large continuing, or non-continuing firms. Non-continuing firms refer to establishments that either exited or entered

[33] Low corresponds to the first two quintiles of the EFE distribution, mid corresponds to the third quintile, and high corresponds to the fourth and fifth quintiles.

the sample from the initial period to the lead period. This category is meant to capture the net effect of new firms. However, this category includes establishments that change ownership or simply were not included in the 2% sample in both periods.[34] Although some new firms may be highly productivity, on average I expect this group to be of lower productivity. Consistent with this view, non-continuing firms have lower average EFEs. Large firms are establishments which where in the top half of the firm size distribution in the initial period and appear in both periods. In terms of employment, the size bins initially have equal employment. Large firms are more likely to survive, however, and as a result make up a larger share of the employment distribution.

Table 1.7: Decomposition of Export Sorting into Worker Flows by Industry and Establishment Fixed Effect

	I. Share of Change in Sorting through Exports by Industry & EFE Distribution						*II. Initial Employment Shares by Industry & EFE Distribution*					
	Manufacturing			Non-Manufacturing			Manufacturing			Non-Manufacturing		
	Low	Mid	High	Low	Mid	High	Low	Mid	High	Low	Mid	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Labor market entry	2.9	-3.8	3.2	32.8	1.1	11.8	2.7	2.9	8.1	10.6	3.7	5.5
Between-LLM job-to-job	*0.2*	*-0.5*	*8.7*	*13.9*	*0.1*	*-5.8*	*1.3*	*1.2*	*3.0*	*5.3*	*1.8*	*2.9*
Within-LLM job-to-job	*1.0*	*-0.5*	*2.6*	*-6.3*	*-0.4*	*3.6*	*1.4*	*1.5*	*3.7*	*4.8*	*1.6*	*2.1*
Job-to-job	1.3	-1.0	11.3	7.7	-0.3	-2.2	2.8	2.7	6.7	10.0	3.4	5.0
Other to emp.	*1.3*	*-0.5*	*7.2*	*9.5*	*-0.1*	*-7.5*	*0.9*	*0.7*	*1.8*	*3.8*	*1.1*	*1.6*
Unemp. to emp.	*-0.2*	*-0.4*	*-1.5*	*1.2*	*0.2*	*1.0*	*0.4*	*0.4*	*1.0*	*1.2*	*0.4*	*0.6*
Nonemployment to emp.	1.1	-0.9	5.7	10.7	0.2	-6.5	1.3	1.1	2.8	5.1	1.5	2.3
Job Stayers	-0.2	1.9	29.1	-3.3	1.1	-3.4	1.8	2.5	7.2	5.5	2.1	2.7
Industry total	5.1	-3.8	49.2	47.9	2.0	-0.3	8.6	9.2	24.9	31.1	10.7	15.5

Notes: Estimates of contributions of worker flows to export-induced LLM sorting based on the methology of Section 1.6.1. Panel I presents the contribution of a given worker flow as a share of the total export-induced change in sorting. Panel II presents the initial share of a given worker flow relative to total LLM employment. "Low" refers to employment in the first and second quintiles of the marginal EFE distribution. "Mid" refers to the third quintle. "High" refers to the fourth and fifth quintiles.

[34] Using the 100% BHP sample, Hethey and Schmieder (2010) find that around 20% of employment of entering and exiting establishments is due to establishment ID changes, spin-offs, or takeovers.

Effects on the Manufacturing Sector

Table 1.7 presents the worker flow decomposition results by industry across the marginal EFE distribution. I begin with a discussion of the results in the manufacturing sector. The most significant worker flow of the manufacturing sector is amplification. Column (3) makes clear that amplification affects the top end of the firm effect distribution. In fact, the effects of amplification are fully accounted for by the changes in employment of firms in the top two quintiles of the EFE distribution. This confirms the result that amplification produces more employment of high-WFE workers in high-EFE firms.

The other source of increasing manufacturing sorting comes from between-LLM reallocation. Similar to amplification, job-to-job transitions affect sorting at the top end of the joint distribution as high-EFE firms account for 8.7% of a total 8.4%. In response to export-induced demand shocks, top manufacturing firms engage in a modest reallocation of their workforce towards high-WFE workers. Other to employment movements are also concentrated at the top end of the joint distribution and show a similarity to between-LLM reallocation. For instance, for both flows the dominant source of sorting is employment changes at high-EFE firms. In contrast, low-EFE firms make the largest relative contribution in terms of unemployment flows. This pattern is repeated in reverse for the non-manufacturing sector. This suggests that the reallocation component of other to employment transitions is more important that the out of the labor force component which I expect to behave similarly to unemployment transitions.

Panel II shows that employment in the manufacturing sector is initially more concentrated in the upper end of the firm fixed effect distribution. For instance, employment in high-EFE firms accounts for 24.9% of initial LLM employment and 58.3% of initial manufacturing employment. However, this over-representation of employment at the top end cannot entirely explain the sorting results. Although 58.3% of initial manufacturing employment is concentrated in high EFE firms, changes in employment in these firms account for 97.4% of the total sorting effect of the manufacturing sector.[35]

Table 1.8 presents the results of the decomposition by industry and firm size. Panel I shows that employment changes of large firms are the dominant source of sorting. Large firms fully account for amplification and account for a majority of both between-LLM reallocation and other to employment flows. Panel II shows that large, surviving manufacturing firms account for 20.6% of total LLM employment and 46.9% of manufacturing employment. Still, they contribute a disproportionate share to sorting–accounting for 39.9% of the total sorting effect and 79.3% of the total manufacturing effect. Interpreting firm size as a proxy for firm productivity, the results of Table 1.8 are consistent with both the results across the EFE distribution of Table 1.7 and some general predictions of trade theory. Large and high-EFE manufacturing firms contribute the most to increasing sorting and they do this by increasing their share of high-WFE workers.

Although these results are broadly consistent with some predictions of trade theory, the majority of the sorting effect works through job stayers. For instance, large firms contribute 30.5% of the export sorting effect through job stayers and 7.3% through reallocation. High-EFE firms contribute 29.1% through job stayers and 11.3% through reallocation. Therefore, most of the demand shock passes through into the wages of current workers rather than towards reallocation of new, high-WFE workers.

Effects on the Non-Manufacturing Sector

Section 1.7.2 shows that the most important worker flow of the non-manufacturing sector is net labor market entry. In fact, there is a relative inflow of entrants to non-manufacturing despite the fact that the demand shock targets the manufacturing sector. I turn to a discussion of the results of the sorting effect of exports on changes in non-manufacturing employment across the firm-type

[35] The final row shows that 49.2% of the total 50.5% effect on sorting due to manufacturing is due to changes in employment at the top end of the EFE distribution.

Table 1.8: Decomposition of Export Sorting into Worker Flows by Industry and Firm Size

	I. Share of Change in Sorting through Exports by Industry & Firm Size						*II. Initial Employment Shares by Industry & Firm Size*					
	Manufacturing			Non-Mnfctr			Manufacturing			Non-Mnfctr		
	NC	Sml	Lrg	NC	Sml	Lrg	NC	Sml	Lrg	NC	Sml	Lrg
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Labor market entry	5.4	2.2	-5.2	28.8	6.6	10.2	4.4	3.3	6.0	10.0	2.6	7.2
Between-LLM job-to-job	*2.5*	*1.1*	*4.8*	*9.7*	*1.3*	*-2.7*	*2.3*	*1.4*	*1.8*	*5.7*	*1.6*	*2.6*
Within-LLM job-to-job	*0.9*	*-0.5*	*2.6*	*1.3*	*1.5*	*3.7*	*3.1*	*1.1*	*3.6*	*5.0*	*0.1*	*-0.1*
Job-to-job	3.4	0.6	7.3	11.0	2.7	1.0	5.4	2.5	5.4	10.7	1.7	2.5
Other to emp.	*0.5*	*-1.2*	*8.7*	*2.6*	*1.7*	*-2.4*	*1.2*	*1.0*	*1.3*	*3.4*	*1.3*	*1.8*
Unemp. to emp.	*0.6*	*-1.3*	*-1.4*	*0.4*	*0.8*	*1.3*	*0.6*	*0.5*	*0.8*	*1.2*	*0.4*	*0.7*
Nonemployment to emp.	1.1	-2.5	7.3	2.9	2.5	-1.1	1.8	1.4	2.1	4.6	1.8	2.4
Job Stayers	0.0	0.2	30.5	0.0	-3.4	-2.2	0.0	4.3	7.2	0.0	3.5	6.7
Industry total	9.9	0.5	39.9	42.8	8.4	7.9	11.7	11.6	20.6	25.3	9.6	18.9

Notes: Estimates of contributions of worker flows to export-induced LLM sorting based on the methology of Section 1.6.1. Panel I presents the contribution of a given worker flow as a share of the total export-induced change in sorting. Panel II presents the initial share of a given worker flow relative to total LLM employment. “NC” refers to employment in non-continuing firms. “Sml” refers to employment in small, surviving firms. “Lrg” refers to employment in large, surviving firms.

distribution to clarify the nature of these flows from labor market entry.

Panel I of Table 1.7 reports the contributions of different worker flows to export-induced sorting across the EFE distribution. Net labor market entry affects sorting through both low- and high-EFE firms. The largest contributors, however, are low EFE firms with a 32.8% contribution which comprises 71.8% of the total contribution of non-manufacturing firms through labor market entry. Although low-EFE firms comprise large share of non-manufacturing employment, their sorting contribution is roughly 50% greater than their employment share. The large contribution of low-EFE firms implies that flows of low-WFE workers to low-EFE firms contribute significantly to increased LLM sorting.

This message is confirmed when looking at the relative contribution of labor market entry by firm size in Table 1.8. Non-continuing firms, either exiting in the lag period or exiting in the lead period, constitute the largest share of the effect of net labor market entry–accounting for 28.8% of the total effect and 62.6% of the non-manufacturing effect. Non-continuing firms comprise 50.5% of employment by labor market entrants and exiters, but yet, produce a disportionate effect on sorting by accounting for 62.6% of the total effect. These results are, therefore, consistent with the interpretation that demand shocks in the manufacturing sector lead to large relative entry of low-WFE workers to new, low-EFE firms in the non-manufacturing sector.

Non-manufacturing sector reallocation also contributes to sorting. In contrast to manufacturing sector reallocation, low-EFE and non-continuing firms fully account for the sorting effects of non-manufacturing reallocation. This provides evidence that the composition of regional entrants to the non-manufacturing sectors shifts toward low-WFE workers.

1.7.4 Implications of Sorting at Labor Market Entry

An important result has been to show that export shocks to the manufacturing sector induce relative movements of entrants away from manufacturing and toward non-manufacturing jobs. These entrants tend to have low WFE's and work for low-EFE firms. On its face this is a counterintuitive result as the manufacturing industry experienced a positive shock. In order to make some sense of this result I will investigate two potential hypotheses.

First, an increase in manufacturing demand may have spillover effects into other industries. Specifically, the growth of the manufacturing sector may lead to an increase in demand for manufacturing inputs. An example of an industry supplying inputs is the business service industry. Business service firms include food, janitorial, and security services.

To investigate this hypothesis I estimate the effect of export shocks on growth in business service industries. I estimate a positive coefficient of 1.201 but it is insignificant with a p-value around 0.12. This null result is consistent with the result on aggregate employment changes in the non-manufacturing industry. Although the export coefficient for change in total non-manufacturing employment is positive at 0.55 log points, it is insignificant with a p-value of 0.22. Therefore, there does not appear to be large spillover demand effects into the non-maunfacturing industry.

A second hypothesis is that firms react to demand shocks by not only increasing employment and output, but by upgrading their technology.[36] Export markets offer the opportunity to increase the scale of operations. An increase in scale induces capital investments that require large fixed costs. Technology on the other hand is ofter viewed as complementarity to a skilled workforce. Therefore, investments may reduce the relative demand for low skill workers. Previous empirical work supports the idea that trade induces firm to upgrade their technology. Both Lileeva and Trefler (2010) and Bustos (2011) find evidence that trade liberalization leads to technological investment in Canada and Argentina, respectively.

In order to investigate this hypothesis, I utilize a data set of employment histories, the LIAB, which can be merge with an establishment-based panel survey, the Establishment History Panel (BHP). The survey provides responses to a variety of questions about the workforce composition

[36]This mechanism is modeling in a trade context by Sampson (2014).

and operations of the establishment. Importantly, it asks respondents to list the total value of all investments undertaken in the previous year.[37] Therefore, the LIAB provides an account of the evolution of firm investment over time.

I replicate the trade shock estimation strategy of equation (1.15) with the data from the LIAB. However, the sample size of the establishment panels is significantly smaller than the SIAB. As the survey is available only from 1990 onwards, I am unable to construct the first estimation interval from 1985 to 1991. The survey sample has expanded through time, so there are more firms in more recent years. In the end I am left with 218 and 648 establishments in the 1990-1996 and 2003 to 2009 intervals, respectively. The sample limitations constrain the regression specification such that I only include state fixed effects and cannot estimate the two-stacked-differences structure.

Despite the sample limitations, I find a significant effect of export shocks on investment in manufacturing firms but not in non-manufacturing firms. I estimate a coefficient on export exposure of 1.107 with a standard error of 0.425. The coefficient is significant at the 0.01% level with a p-value of 0.009.[38] This result is consistent with a story in which manufacturing firms upgrade their technology and reduce their demand for low-skill labor. Young entrants are then left to find jobs in less well-compensated industries such as the retail trade and service sectors. This story fits with the long-term features of the manufacturing industry. Employment in manufacturing, even in Germany, declined significantly from the 1980's through the 2000's. However, manufacturing output has continued to rise. Therefore, capital increasingly plays a larger role in the production of manufacturing goods and increases in demand appear to accelerate this process.

[37] They also asks respondents to evaluate the technical status of their equipment. Respondents are given a choice of stating the state of their equipment as either "obsolete", "rather obsolete", "medium", "rather state-of-the-art", or "state-of-the-art". Given the discrete nature of the responses and the fact that the majority of establishments pick "rather state-of-the-art", this variable may be a less objective and reliable measure.

[38] I fail to replicate my results in the SIAB on employment and wages. The coefficients are positive, but insignificant. I suspect the insignificance of these results are due to the small sample sizes. There are far too few firms in each LLM to construct an accurate measure of within-LLM sorting to attempt to replicate the sorting resutls. Results of these regressions are reported in the appendix.

1.8 Conclusion

Although labor market sorting is a significant contributor to inequality, the channels and sources of its rise remain unclear. I combine a novel worker flow decomposition method with exogenous trade shocks to understand how and why sorting is rising. My main finding is that young workers are becoming more sorted at labor market entry over time. By performing a decomposition of exogenous, export-induced worker flows, I confirm this descriptive statistic and alleviate concerns over changes in worker composition. For a given distribution of skill types, entrants are becoming more sorted in recent decades, i.e. high-wage entrants are matching with high-wage firms.

The finding of increased sorting at labor market entry has important implications for inequality. As high-wage workers are sorted into high-wage firm at earlier stages in their careers, the effects on lifetime inequality will be greater as high-wage workers earn firm premiums for longer durations. Furthermore, an increase of sorting at entry suggests that the factors that generate individuals' initial productivity, such as education and childhood environment, are important determinants of lifetime sorting and inequality.

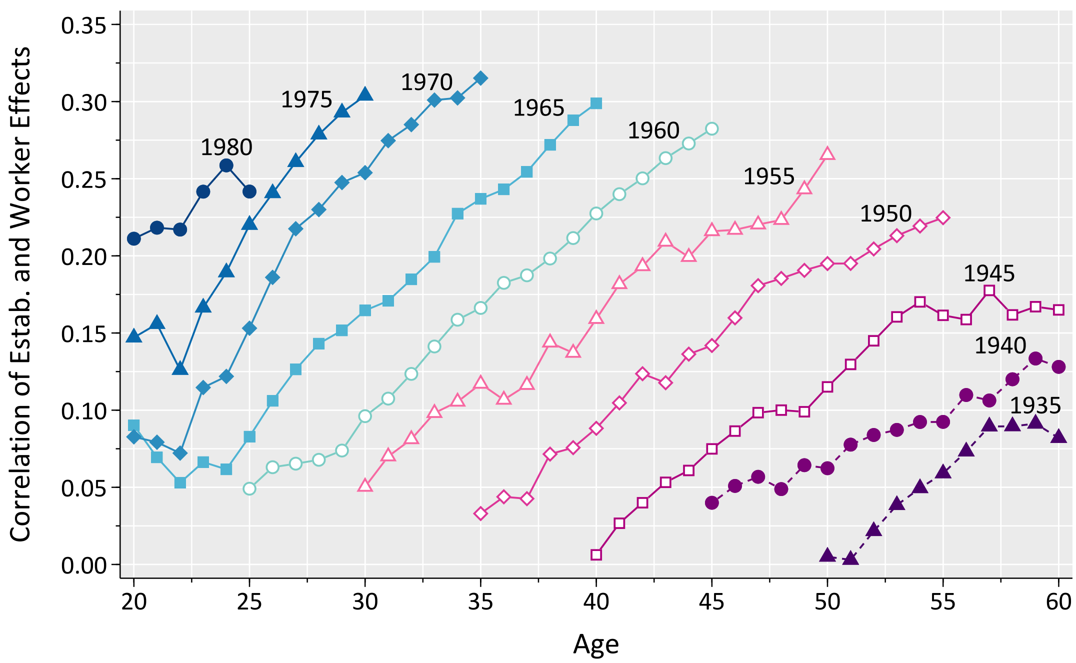
In addition, these results highlight demand shocks as a potentially important source of rising sorting. Trade from Eastern Europe and China alone can account for 14% of the total rise in sorting in West German over the period 1985 to 2009. Given that the export-induced sorting flows are similar to the aggregate sorting flows and trade with the East accounts for small portion of international and domestic trade, increases in demand can potentially explain a large share of the total rise in West German sorting. This result is consistent with a story of rising sorting due to complementarities between firm technology and worker skill which are amplified by demand shocks.

By analyzing export-induced worker flows, we gain a better understanding of the impact of exports on sorting. I find that shocks to the manufacturing industry have large effects on labor market entry in non-manufacturing sectors. Specifically, low-wage workers tend to enter to low-wage firms. To explain this result I hypothesize that manufacturing firms increase investment to take advantage of increases in demand. As a result, low-skill workers have few employment opportunities and, therefore, enter industries that pay lower firm premiums. I find support for this hypothesis by estimating a significant effect of export shocks on investment.

1.9 Appendix: Descriptive Patterns of Sorting Over the Life-Cycle and Over Time

The purpose of this section is to identify potentially important factors for the rise of sorting with descriptive evidence. This evidence provides a basic understanding of how the sorting process evolves over the career and how these processes have been changing over time. To provide an organizational structure, I describe the potential factors in terms of age, year, and cohort effects.

Figure 1.2: Sorting Over the Life-Cycle by Year-of-Birth Cohort



Notes: YOB cohorts labeled within the plot region. Each data point represents the average value of the correlation of establishment and worker fixed effects across five YOB cohorts. For example, 1980 represents an average across all years from 1980 to 1984.

Figure 1.2 plots sorting of worker and firm types by year-of-birth (YOB) cohort and age. Consistent with Section 1.3, sorting is defined as the correlation between establishment and worker fixed effects. A notable distinction here, however, is that this measure is computed *within* YOB and age groups. Total labor market sorting is a function of both within- and between-group effect.

This analysis, therefore, focuses on one component of total sorting.[39] Each data point in the figure represents an average measure of sorting across five separate YOB cohorts. The label for each line corresponds to the earliest year of a five-year group. For example, the line labeled 1935 represents average sorting for YOB cohorts 1935 through 1939 at each specified age.

Figure 1.2 presents two key facts. First, within a given cohort, sorting is rising with age. This fact is represented in the positive slope of each of the YOB cohort lines. Second, sorting is higher in younger cohorts. Indeed, at any given age, sorting is always higher for the younger than the older cohort.[40]

In terms of age, year, and cohort effects, there are two stylized hypotheses that can explain these facts. The first hypothesis is that age effects are positive and constant, year effects are constant, and cohort effects are increasing. The second hypothesis is that age effects are constant, year effects are increasing, and cohort effects are constant.[41] Throughout the course of this section, I argue that the first hypothesis fits the data best and, therefore, the rise is sorting is the result of increasing cohort effects.

The fact that age, year, and cohort effects cannot be jointing identified is a classic problem without a ready solution. Researchers need to make assumptions about some features of these effects in order to jointly identify them.[42] Given that the life-cycle pattern of sorting is poorly understood, I refrain from using one of these methods. In the following discussion, I provide an

[39] Section 1.12 suggests that within-group effects may be the more important component of the rise in sorting. The results indicate that most of the change in sorting is due to within-group changes in the joint distribution of establishment and worker effects. However, the "groups" in this section include worker flow states, firm size, and industry.

[40] The single expection is at age 20 YOB cohort 1965 is more sorted than YOB cohort 1970.

[41] A third stylized hypothesis woudl be that year and cohort effects are constant and age effects are increaseing over time–creating a steeper slope in the age profile of sorting. However, this hypothesis seems difficult to reconcile with the paterns in Figure 1.2.

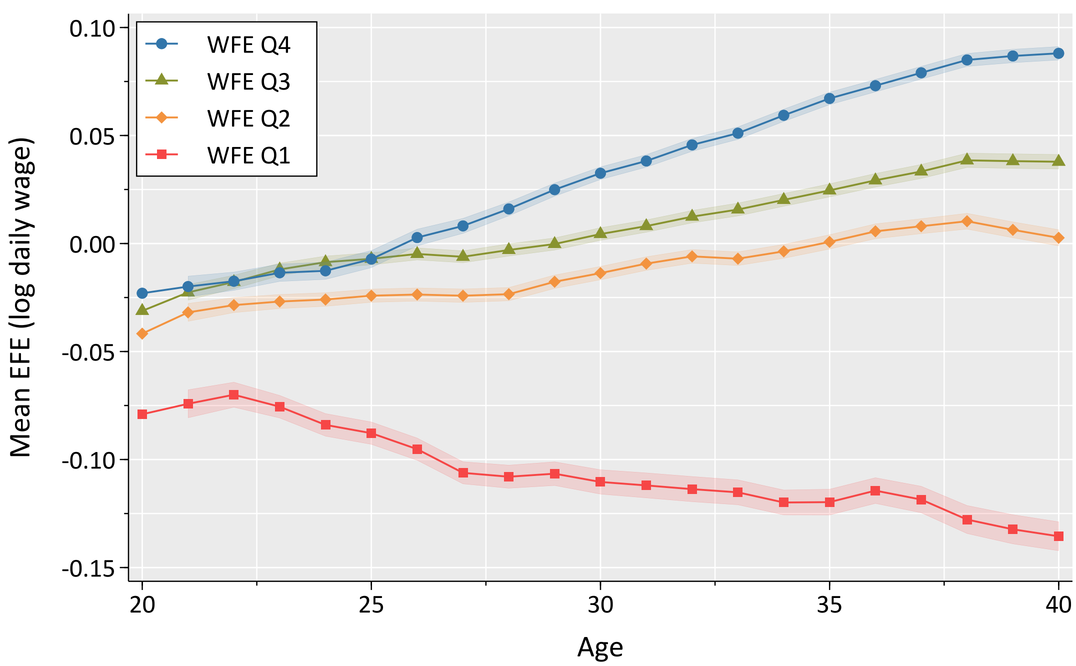
[42] See Hall (1968), Deaton (1997), Card and Lemieux (2001), Heckman et al. (1998), and Lagakos et al. (2016).

explanation of the two key facts of Figure 1.2 using patterns in the data. Although the evidence presented in this section is descriptive, it helps build intuition to understand the causal evidence presented in Section 1.5.

1.9.1 Fact 1: Why does sorting rise with age?

Before considering changes in sorting over time, first consider the growth in sorting over the life-cycle. For expositional purposes, I focus on the 1965-1969 YOB cohort. For this cohort, my sample covers 20 years of observations from age 20 to 40. In Section 1.9.2, I turn to the question of whether sorting patterns are stable over time.

Figure 1.3: Average EFE Over the Life-Cycle by WFE Quartile for YOB Cohort 1965-1969



Notes: Each age t on the x-axis represents the transition between age $t-1$ and t.

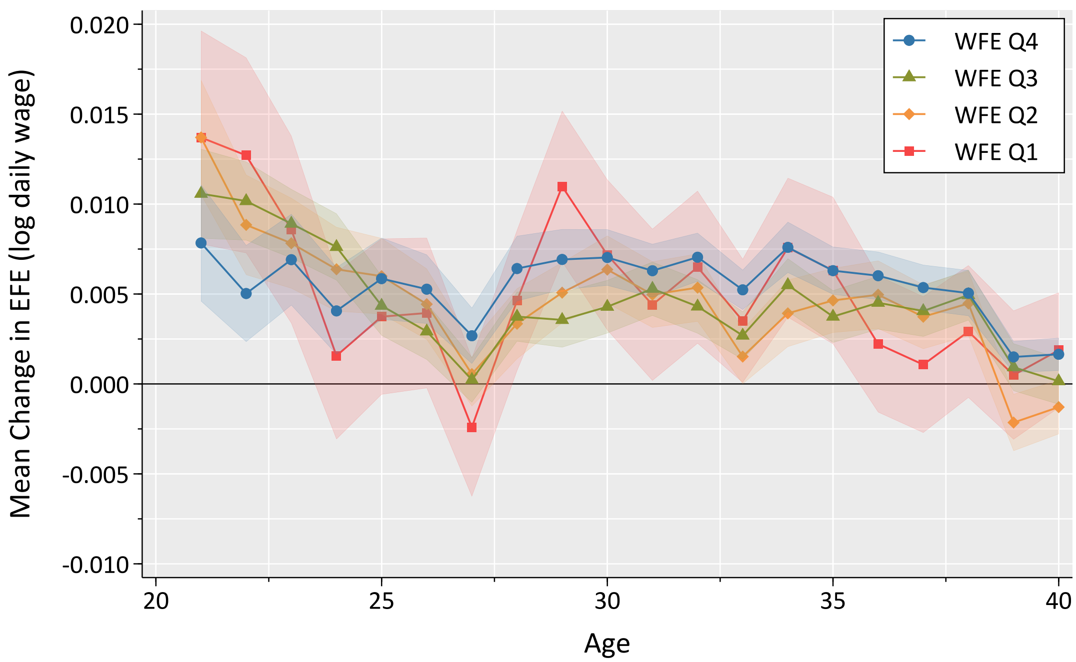
Figure 1.3 provides an alternative representation of the rise in sorting over the life-cycle. Each individual is classified into one of four worker fixed effect (WFE) quartiles between the ages of 20 and 40. These quartiles are computed at the worker level, therefore, at any given age there may be more individuals in a given quartile if they are more likely to be employed. Each line plots the average establishment fixed effect (EFE) for each quartile at each age.

The figure shows a few key results. First, consider the early career stages from ages 20 to 25. Initially at age 20, the rank ordering of average EFE is consistent with the rank ordering of WFE quartile, with WFE Q4 earning the highest firm premium and WFE Q1 the lowest. In the next five years, the average EFE for Q2 and Q3 rises steadily, while the average EFE for Q4 rises only moderately, and the average EFE for Q4 falls slightly after an initial increase. Given that EFEs of WFE Q2 and Q3 workers rise faster than WFE Q4 workers, the effect on sorting is ambiguous. Indeed, Figure 1.2 shows that sorting is roughly constant during this period (initially falling and subsequently rising).

After the age of 25, however, we see a clear trend toward greater dispersion in EFEs across WFE quartiles. The average EFE of WFE Q4 rises rapidly, WFE Q2 and Q3 resume their steady march upward, and the averge EFE of WFE Q1 workers continually declines. Over the course of 20 years, the difference between the average EFE in WFE Q4 and WFE Q1 grows from about 10 log points at age 20 to 22.5 log points at age 40. This in turn leads to a sharp increase in within-cohort sorting, as shown in Figure 1.2. To understand life-cycle sorting, two questions become apparent. First, why does the average EFE of WFE Q1 workers fall? Second, why does the average EFE of WFE Q4 workers rise faster than WFE Q2 and Q3 workers?

To address these questions, I deconstruct the changes in EFEs over time into the effects of two types of transitions: employment-to-employment and employment-to-nonemployment. This classification helps us to understand whether the differential growth in EFEs is caused by the return to staying employed or the cost of incurring a nonemployment transition. In investigating these transitions, I report statistics on the differential probabilities and returns to each transition across workers types.

Figure 1.4: Change in EFE Given Stay Employed Over the Life-Cycle by WFE Quartile for YOB Cohort 1965-1969



Notes: Each age t on the x-axis represents the transition between age $t-1$ and t.

The literature typically focuses on the role of job-to-job transitions to explain sorting (e.g. Hagedorn et al. 2017, Lopes de Melo 2018). Therefore, one natural story for increasing dispersion in EFEs is that high WFE workers move to higher wage firms over time while on the job. To investigate this channel, Figure 1.4 plots the average change in EFE by each WFE quartile for one-year transitions. As the figure includes all employment-to-employment transitions it accounts for both changes in the EFE due to job changes and increases in an estlabishment's EFE over time.[43]

The main result of Figure 1.4 is that all workers tend to incur similar growth in their EFEs while employed–and this is particularly true after age 25. The shaded regions of the figure represent 95% confidence intervals for each quartile. As all the confidence intervals overlap past age 25, we are unlikely to reject the null hypothesis that the gains in EFEs while staying employed are the same across worker types. Therefore, perhaps surprisingly, differential returns to job-to-job transitions are unlikely to account for the life-cycle profile in sorting.

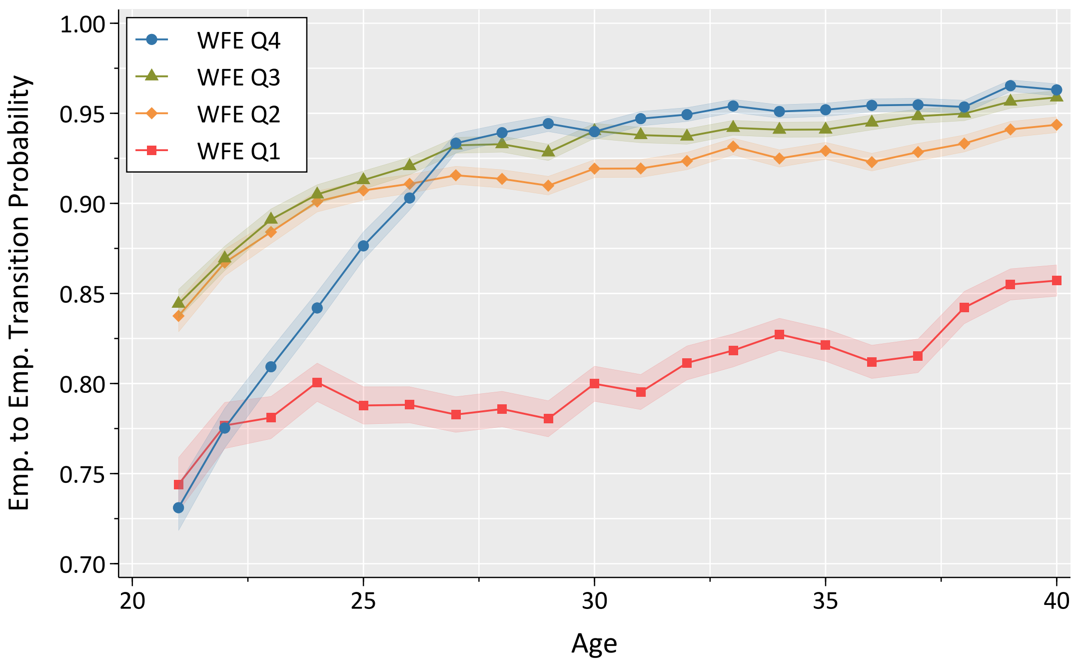
Another interesting result is that until about age 38 most workers tend to improve their EFE with experience. In order words, early career workers tend to slowly transition to higher paying firms. This is partly the result of the fact that workers tend to transition to higher paying firms and partly that surviving firms tend to increase their firm wage premiums.

Another possible explanation for the growth in sorting over the life-cycle is that different worker types have different probabilities of staying employed. If workers face a loss in their EFE following a nonemployment transition, then a higher incidence of these transitions will lead to slower growth in EFEs.

Figure 1.5 plots the annual probability of staying employed by WFE quartile and age. A few features are of note. First, there is a steep rise in the probability of staying employed for the top WFE quartile of workers from ages 21 to 27. This feature suggests that transitions between the

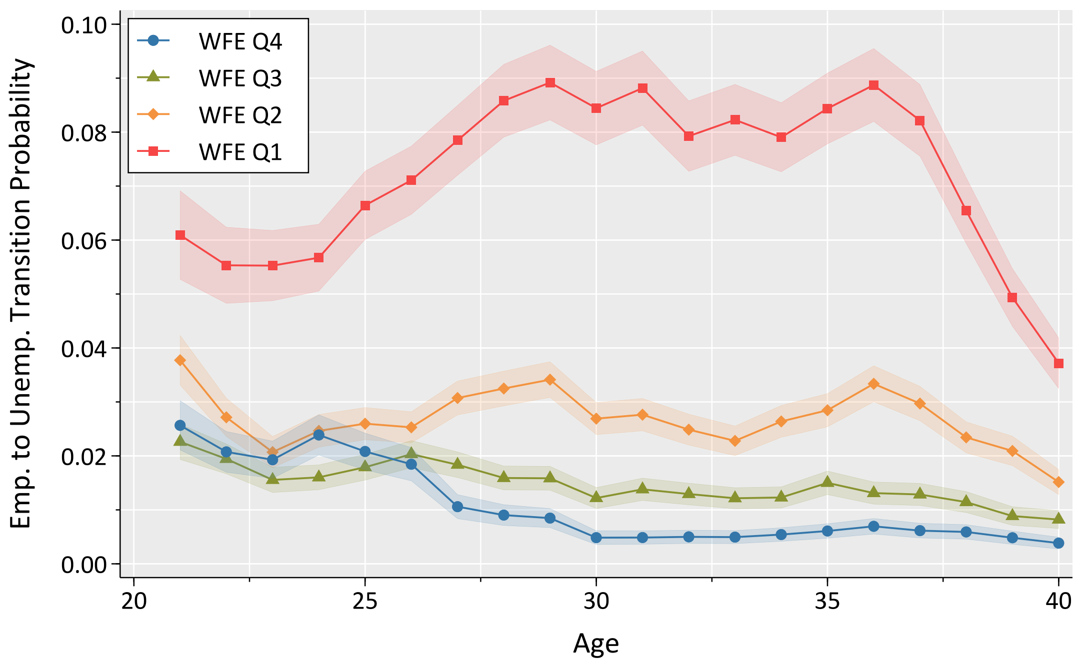
[43] EFEs are computed for the same firm in multiple periods and, therefore, may grow over time. More detailed figures of the effects of changes in EFEs due to job switching and job staying can be produced upon request.

Figure 1.5: Probability to Stay Employed Over the Life-Cycle by WFE Quartile for YOB Cohort 1965-1969



Notes: Each age t on the x-axis represents the transition between age $t-1$ and t.

Figure 1.6: Probability to Transition to Unemployment Over the Life-Cycle by WFE Quartile for YOB Cohort 1965-1969



Notes: Each age t on the x-axis represents the transition between age $t-1$ and t.

labor force and schooling last until the late 20s WFE Q4 workers. Also it suggests that high-wage workers are indeed high-skill workers. After the age of 28, however, WFE Q4 workers are the most likely to stay employed. Although the differences in the probability of staying employed between WFE Q3 and Q4 are quite small, the standard errors are small enough to, in many cases, produce statistically significant differences.

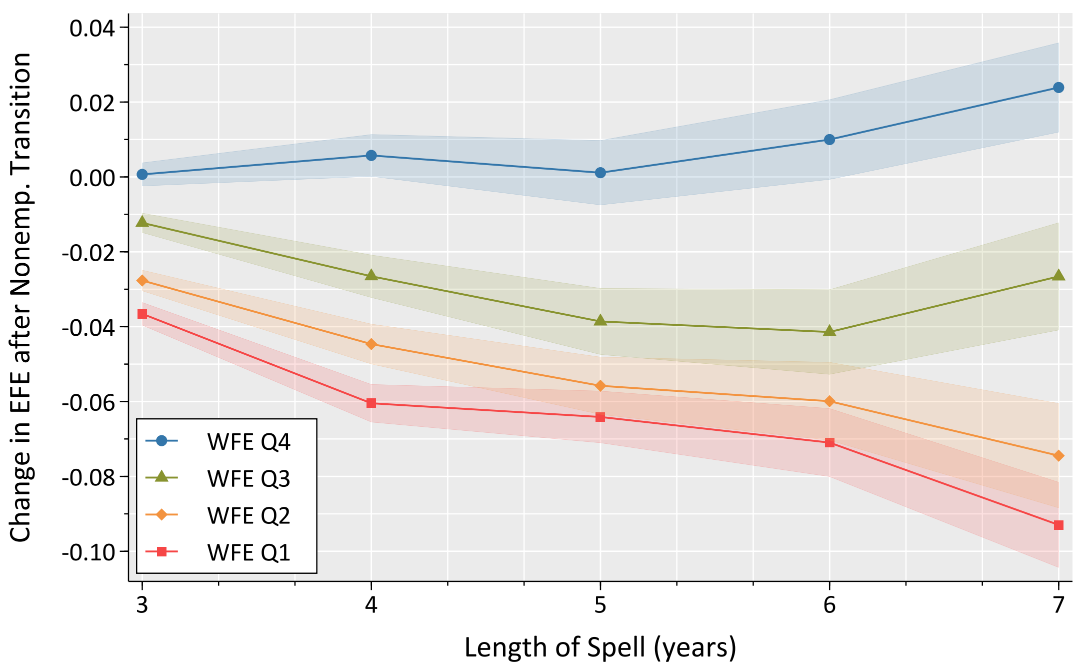
Another striking feature of Figure 1.5 is that WFE Q1 workers face a low probability of staying employed. From ages 20 to 40, WFE Q1 workers are consistently over 10% more likely to face a nonemployment transition. Given the stability of this disparity over the life-cycle, it is unlikely that these differences reflect differences in the probability of transitioning to schooling, as we would expect schooling transitions to be concentrated in the early career. In fact, Figure 1.6 shows that WFE Q1 workers face a consistently higher probability of transitioning to unemployment with an average disparity of about 6% between ages 29 and 37.

Given that workers with low WFEs are significantly more likely to face a nonemployment transition, this could account for the negative growth in the EFEs of WFE Q1 workers provided that the cost of a nonempoyment spell is significant. In fact, many search models predict that workers face a risk of falling to the bottom of the job ladder after a nonemployment transition (e.g. Burdett and Mortensen 1998, Delacroix and Shi 2006, Jarosch 2015, and Krolikowski 2017).

Figure 1.7 plots estimates of the cost of nonemployment transitions by comparing the initial EFE before a nonemployment spell with the re-employment EFE. The x-axis reports the spell length of an employment sequence ranging from three to seven years. Using this notation, a three-year spell represents a sequence of employment states such that the individual is employed in the first year, nonemployed in second year, and then re-employed in the third year. Hence the sequence labeled as a three-year spell represents changes in EFEs across two years.

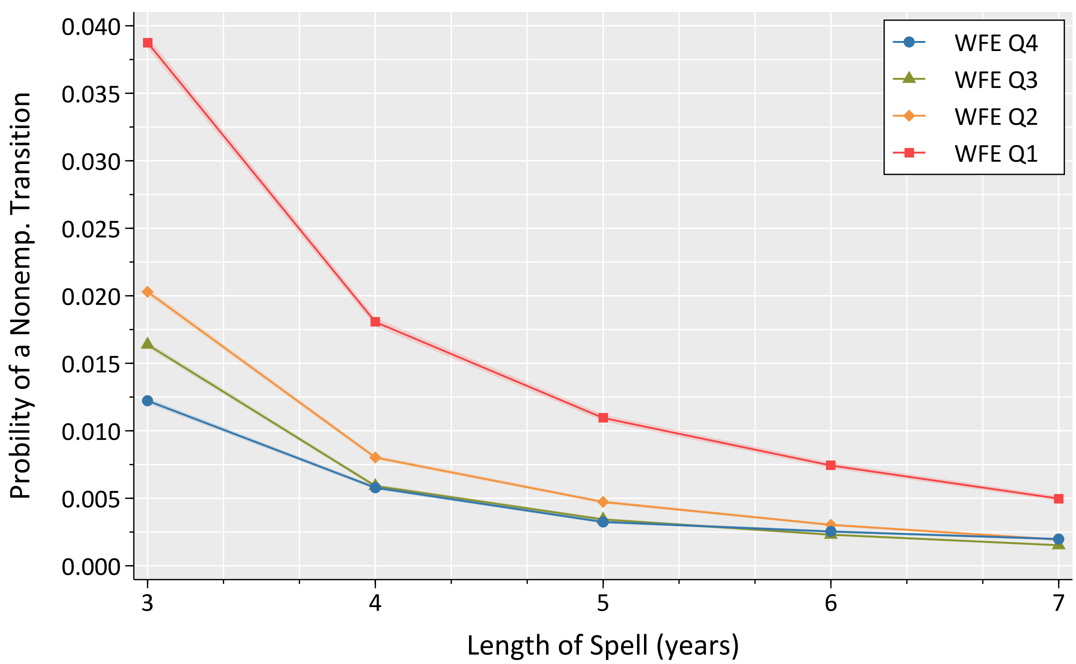
Figure 1.7 shows that for WFE Q1, Q2, and Q3 workers there is a significant cost of incurring a nonemployment spell. Furthermore, the cost of the transition grows with the duration of the

Figure 1.7: Cost of a Nonemployment Transition by Worker Type



Notes: A spell with a nonemployment transition includes a spell of nonemployed bookended by a year of employment. For example, a spell of length 5 denotes the sequence: E, N, N, N, E, where E denotes employment and N denotes nonemployment. Differences in wages for a spell of length t are then computed across $t-1$ years.

Figure 1.8: Probability of Incurring a Nonemployment Transition by Worker Type



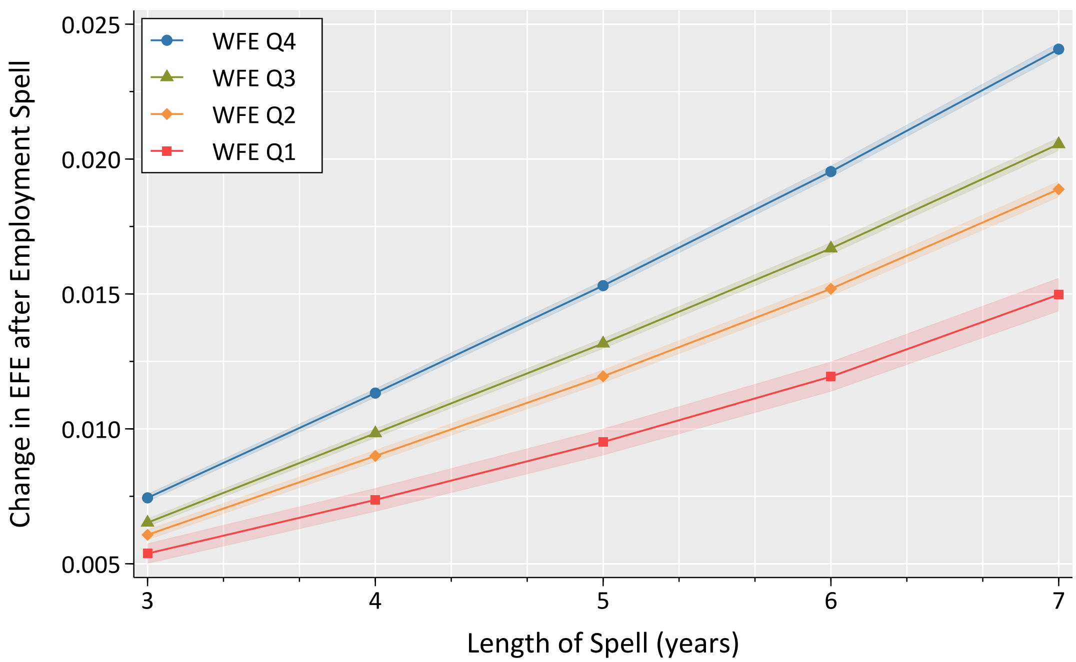
Notes: A spell with a nonemployment transition includes a spell of nonemployed bookended by a year of employment. For example, a spell of length 5 denotes the sequence: E, N, N, N, E, where E denotes employment and N denotes nonemployment.

spell. For instance, Q1 workers initially face an average loss in EFEs of about 4 log points. This loss grows to about an 11 log points in a seven-year spell (five years of nonemployment). On the other hand, WFE Q4 workers do not suffer losses in their EFE when transitioning through nonemployment. These differences further suggest that high-wage workers experience different types of nonemployment spells than low-wage workers. If high WFE workers are more likely to transition through school rather than unemployment, the cost of nonemployment would be expected to be smaller.

Both the high incidence and high cost of nonemployment for WFE Q1 workers suggests that the nonemployment channel is an important source of the decline in EFEs for low wage workers over the life-cycle and, hence, an important factor in the rise of life-cycle sorting. To understand the magnitude of this channel, Figure 1.4 shows that growth in the mean EFE for WFE Q1 workers is -5.65 log points from age 20 to 40 whereas the mean EFE for WFE Q2 workers grows 4.44 log points. Hence the gap in EFEs between WFE Q1 and Q2 grows by 10.09 log points in 20 years. Considering only nonemployment spells of five year or less (total spells of seven years or less), I use the cost of nonemployment spells (reported in Figure 1.7) along with the probability of incurring such a spell (reported in Figure 1.8) to estimate that over 20 years nonemployment spells result in a relative loss in EFEs of 4.45 log points for WFE Q1 relative to WFE Q2. This represents about 45% of the total career mean EFE growth differential of 10.09 log points, but is a lower bound since it does not included nonemployment spells longer than five consecutive years.

On the other hand, Figure 1.9 reports the cumulative growth in EFEs of workers who stay employed by worker type. Even while consistently employed, the dispersion in mean EFEs grows with time. Although these differences are statistically significant, they are small. For example, if I linearly extrapolate the growth in EFE dispersion out to 20 years, the difference in mean EFE growth between WFE Q1 and WFE Q2 amounts to -0.93 log points. This differential represents an estimate of difference in EFE growth if both groups were fully employed over the ages 20 to 40 and, hence, is an upper bound of the effect of on-the-job EFE growth. Therefore, nonemployment transitions appear to be the main determinant of increased sorting at the low end of the WFE distribution.

Figure 1.9: Return to Staying Employed by Worker Type

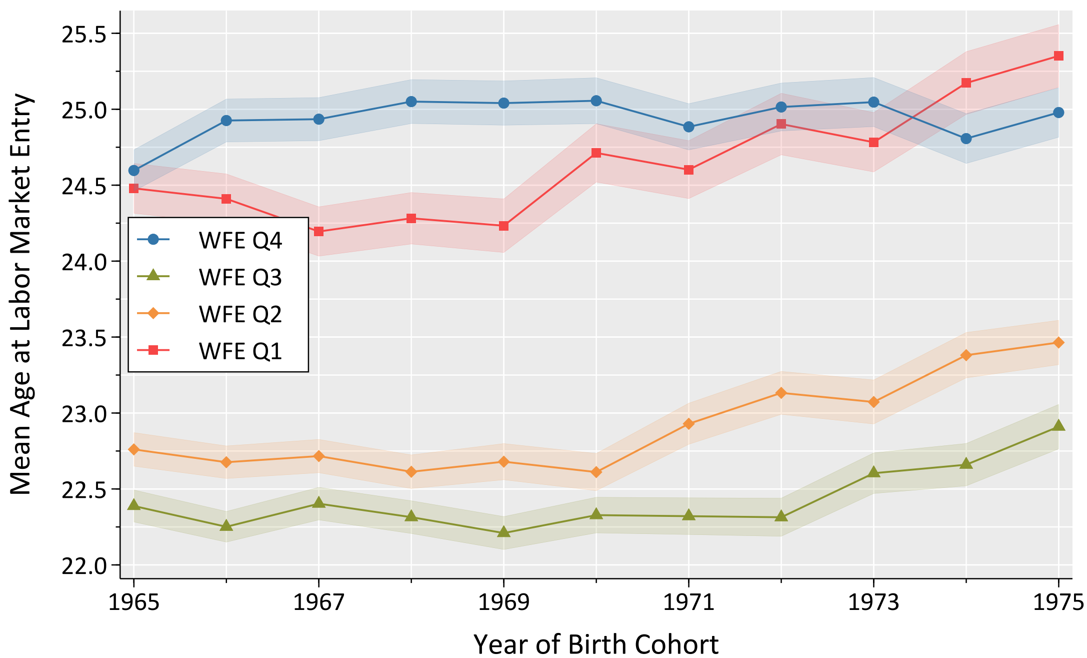


Notes: Differences in wages for a spell of length t are computed across $t-1$ years.

Another factor in rising sorting is the steep rise in EFEs for high-wage workers throughout their careers. Figure 1.5 shows that WFE Q4 workers are relatively less attached to the labor force than WFE Q3 or Q2 workers until the age of 27. Due to their high lifetime earnings potential, this instability is likely the result of transitions in and out of schooling. Therefore, the initial low level and low growth of EFEs in Figure 1.3 likely results from the fact that many of these workers are taking temporary jobs with firms that due not reflect their full earnings potential. The faster growth in the average EFE between ages 25 and 27 likely reflects labor market entry into more permanent career jobs.

Figure 1.10 provides some evidence for the claim that high WFE workers are more likely to permanently enter the labor market at older ages. The figure plots the average age at which individuals begin their first three consecutive years of full-time employment between the ages 20 to 32.[44] Note that WFE Q4 workers consistently enter one to two years later than WFE Q3 and Q2 workers. WFE Q1 workers also enter later, but this likely reflects the fact that WFE Q1 workers are not firmly attached to the labor force at any point. Therefore, a more appropriate definition of labor market entry for Q1 workers may be the first year of full-time employment.

Figure 1.10: Average age of "permanent" labor market entry by WFE quartile and YOB cohort



Notes: "Permanent" labor market entry denotes the first consecutive three-year employment spell of an individual between the ages of 20 and 32.

Given the differential entry patterns between WFE Q4 workers and WFE Q3 workers, selection may play an important role in explaining different life-cycle trends in the early career years

[44] I restrict the age of entry between ages 20 and 32 to make the statistic comparable across YOB cohorts. Since my sample ends in 2009, I observe older cohorts for more years and at older ages. Hence, the sample contains more years for these cohorts to potentially enter the labor market. Late entrants can push the average age of entry up. To make the statistic comparable I condition on an age range over which all cohorts are equally represented in the sample.

between these groups. In any case, most of the differential growth rate in average EFE comes after the age of 27. Thus I focus on differential EFE growth from age 27 to 40 for my analysis of the top end of the worker skill distribution.

There appear to be two reasons for the divergence in EFE growth. Figure 1.5 shows that WFE Q4 workers are more slightly more likely to stay employed at all ages from 27 to 40 with an average differential between WFE Q3 (WFE Q2) workers of 0.9% (2.6%) during these years. Furthermore, Figure 1.6 shows that most of this difference is due to to fact that Q4 workers are likely to experience a transition to unemployment and, hence, less likely to experience a costly nonemployment transition. The average differential in the probability of incurring an unemployment transition between WFE Q4 and WFE Q3 (WFE Q2) workers is 0.7% (2.1%).

Another cause of the higher growth in EFEs for WFE Q4 workers is that they experience slightly higher growth in EFEs while staying employed. In any given year, the difference in earnings growth is insignificant (Figure 1.4), but over time significant differences materialize (Figure 1.9). To put the magnitude of the two channels in perspective. Figure 1.4 implies that the growth in EFEs of WFE Q4 workers was greater than WFE Q3 workers by 3.59 log points from age 27 to 40. The estimate of the differential cost of nonemployment transitions between WFE Q4 and Q3 workers using only nonemployment spells of five years or less for this length period is 0.70 log points (a lower bound). The estimate of the differential return to staying employed between WFE Q4 and Q3 for 13 years is 1.16 log points (an upper bound as the estimate assumes full employment for each group over the period.) Therefore, the two channels are of roughly equal magnitude, suggesting that the movements up the job ladder are more consequential for sorting at the upper end of the worker skill distribution.

In conclusion, an important driver of the life-cycle growth in within-cohort sorting is the differential incidence of nonemployment transitions by different worker types. The lowest WFE quartile workers have a substantially larger chance of incurring a costly nonemployment spell which disrupts their ascent up the job ladder. On the other hand, WFE Q4 workers both have slightly a lower chance of incurring an unemployment spell than WFE Q3 workers and a slightly higher return to staying employed which accumulate over a career to produce dispersion in the average EFE at the top of the WFE distribution.

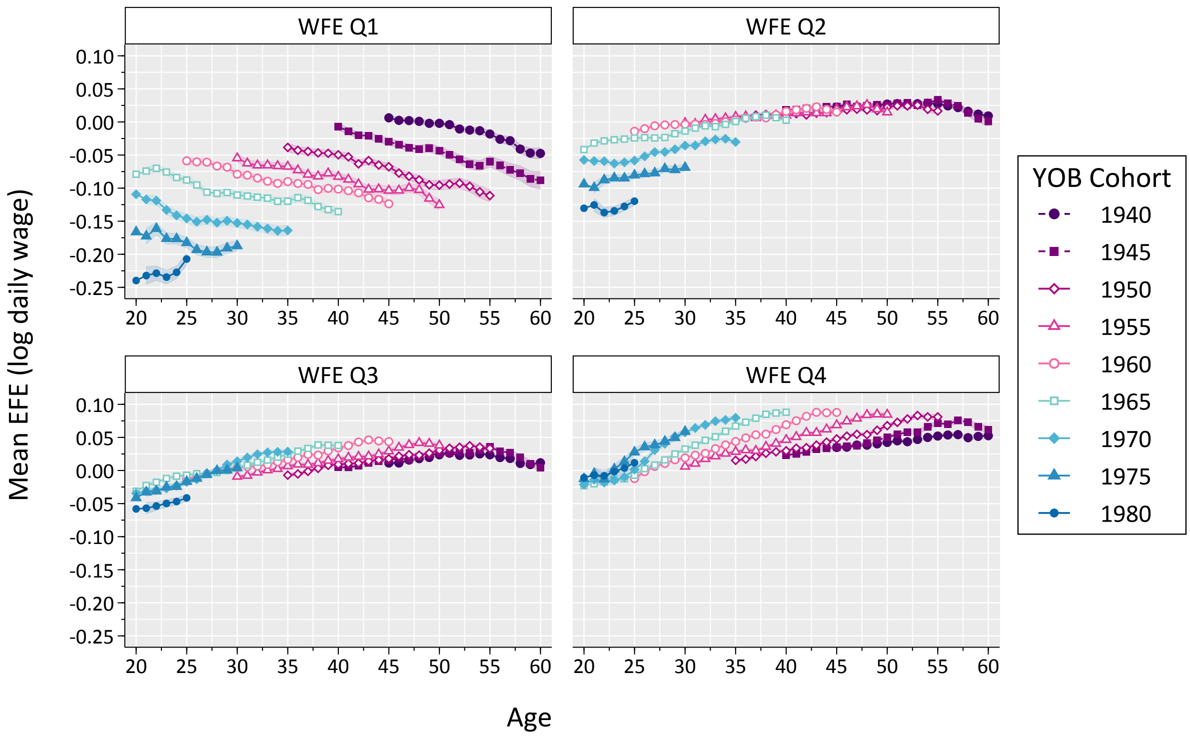
1.9.2 Fact 2: Why is sorting rising over time?

I now turn to an investigation of the changes in sorting patterns over time. Figure 1.11 replicates the same information as Figure 1.3 across different YOB cohorts. For each cohort, the y-axis records the average EFE and the x-axis records age. The fact that EFEs of WFE Q1 workers shift downward over time and EFEs of WFE Q4 workers shift up over time indicates that sorting is increasing over time. In fact, for WFE Q1 workers at age 40 their average EFE declines by about 12.5 log points from YOB cohort 1945 to YOB cohort 1985. On the other hand, for WFE Q4 workers at age 40 their average EFE increases about 6 log points from the 1945 to the 1965 YOB cohort. Thus the total differential between WFE Q1 and Q4 workers widens by about 18.5 log points in 20 years.

Three broad potential mechanisms could be producing this dispersion over time. First, there could be a change in the relative probability of staying employed across worker types. Second, there could be a change in the relative return to staying employed or the relative cost of nonemployment spells across worker types. And finally, there may be a trend towards greater dispersion in the initial average EFE at labor market entry across worker types. Below I argue that the descriptive statistics support the final hypothesis, i.e. growing dispersion in initial EFEs explains most of the rise in sorting over time. Furthermore, I argue that growing dispersion in EFEs at labor market entry is consistent with an increasing cohort effect.

I discuss each hypothesis sequentially. Figure 1.12 shows that probability of staying employed, and hence the probability of transitioning to nonemployment, is roughly stable over time. This

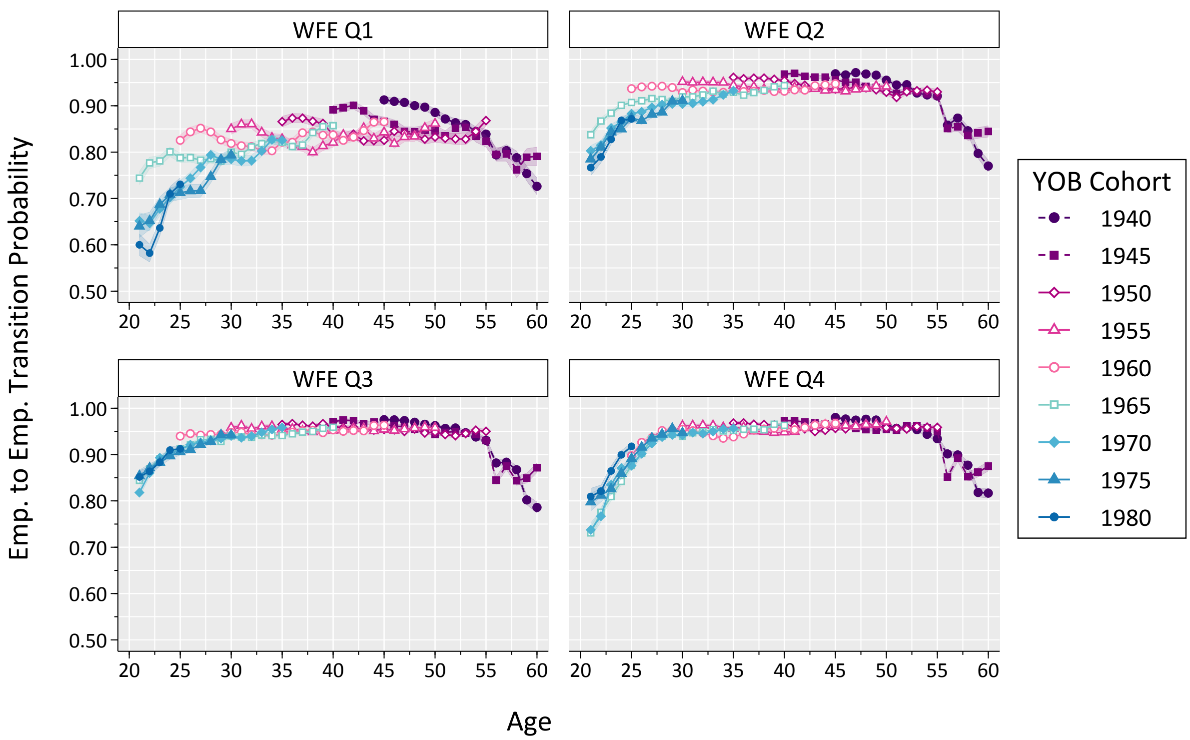
Figure 1.11: Average EFE Over the Life-Cycle by WFE Quartile and Year-of-Birth Cohort



Notes:

conclusion is clear for WFE Q3 and Q4 workers where, despite the fact that the standard errors are very small, in almost all cases with multiple observations at the same age across YOB cohorts, the confidence intervals overlap significantly. The picture is similar for WFE Q1 and Q2 with one exception. It appears that during the period from 1989 through 1993 there was a significant reduction in the probability of staying employed for workers of all ages in WFE Q1. This may suggest that German reunification put pressure on the lower end of the worker skill distribution as the supply of low-skill workers increased due to migration from East Germany and Eastern Europe. Therefore, the fall in employment security for low-wage workers represents one aspect of the rise in sorting, but only for a specific period. This would be captured as a year effect in the age/year/cohort effect framework since the effect is present for all age groups.

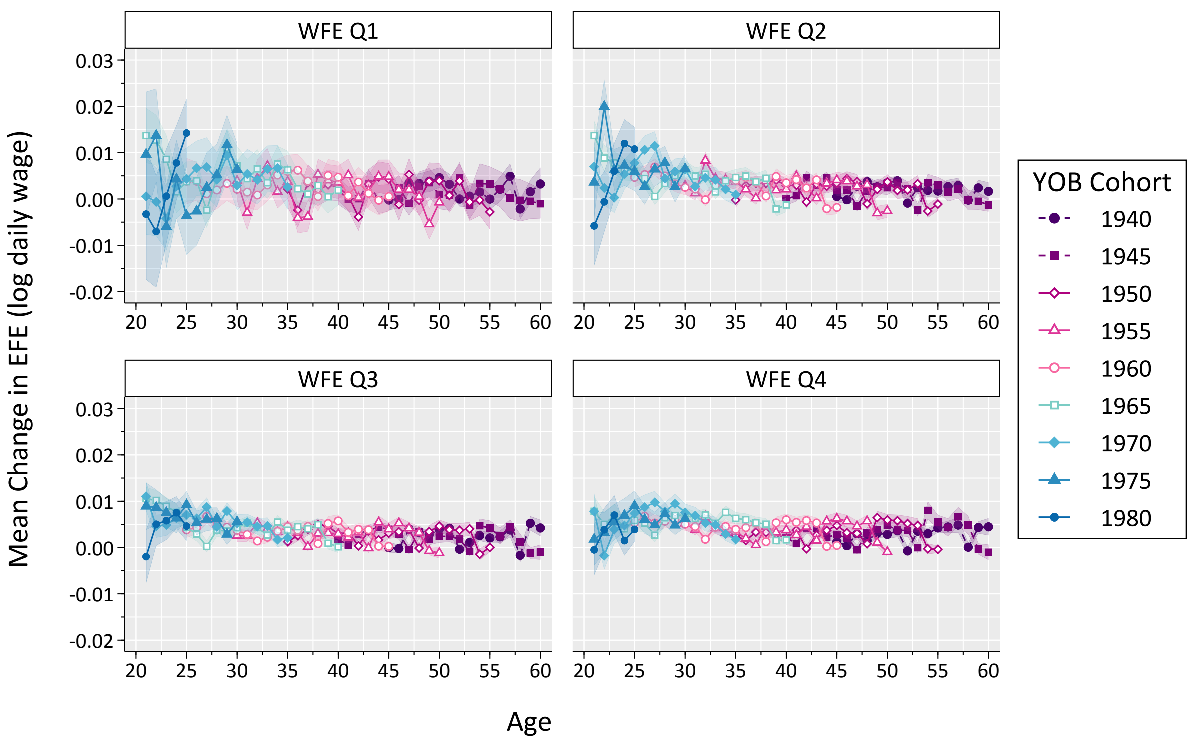
Figure 1.12: Probability Stay Employed Over the Life-Cycle by WFE Quartile for YOB Cohort



Notes:

Figure 1.13 shows that the return to staying employed has been roughly constant over time. As evidence for this conclusion, notice how for each WFE quartile the lines demarcating the return to staying employed lie roughly on top of each other across cohorts with the confidence intervals generally overlapping. This suggests that the return to job-to-job transitions and staying employed in a surviving firm remain roughly constant over time. Although not pictured, the cost of a nonemployment transition also remains roughly constant over time. Therefore, Figures 1.12 and 1.11 show that the life-cycle sorting patterns, in terms of the incidence and return to employment and nonemployment transitions, appear to be relatively stable over time.

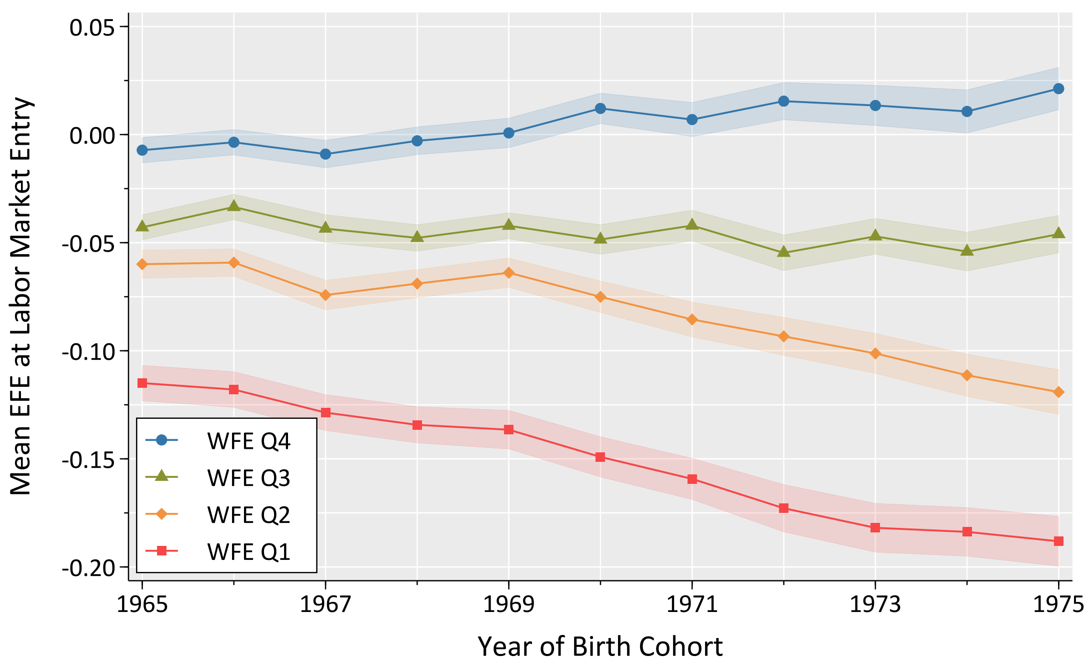
Figure 1.13: Change in EFE Given Stayed Employed Over the Life-Cycle by WFE Quartile and YOB Cohort



Notes:

On the other hand, the dispersion in the average EFE at labor market entry appears to be increasingly substantially. Figure 1.14 plots the average EFE at labor market entry for each WFE quartile by YOB cohort. Since this statistic relies on the ability to observe workers' employment histories between ages 20 and 32, the statistic can only be computed for a limited number of YOB cohorts. Despite short interval of observation, there is strong evidence of increasing dispersion in the average EFE at labor market entry. For instance, the differential between average EFEs at labor market entry for WFE Q4 versus WFE Q1 workers widens from 10.8 log points for YOB cohort 1965 to 20.9 log points for YOB cohort 1975. Therefore, the gap in average EFE at labor market entry grows 10.2 log points in just ten years. Figure 1.11 shows an increase in the dispersion of average EFEs between WFE Q4 and Q1 of about 18.5 log points in the 20 years from YOB cohort 1965 to 1985. A linear extrapolation of the change in average EFE at entry results in a 20.4 log point dispersion–accounting for the full effect.[45]

Figure 1.14: Average EFE by WFE quartile and YOB Cohort



Notes:

[45]With an important caveat being that the calculations of the total change and effect of labor market entry come from different time periods.

The probability of staying employed, the return to staying employed, and the cost of incurring a nonemployment transition are all approximately constant over time. Given that these factors fully characterize career growth in EFEs, the life-cycle pattern of sorting must also be approximately constant over time. Therefore, growth in sorting is not likely to be the result of age or year effects– factors which vary over the life-cycle. On the other hand, the data clearly shows growing dispersion in the initial EFE at labor market entry across WFE types. Thus the growth in sorting can be characterized as a cohort effect, where we see a constant life-cycle profile that is steadily shifting up over time.

1.10 Appendix: Robustness of the Effect of Export Exposure on Sorting

Table 1.9 presents the complete results of the estimation of the effect of trade exposure shocks on local labor market sorting (equation 1.15). Column (1) controls for broad geographic trends with West Germany divided into 74 regions. Columns (2), the main specification, controls for more narrow geographic trends with West Germany divided into 214 regions. The first two rows report the coefficients of the endogenous variables (changes in import and export exposure) while the remaining rows present the coefficients of control variables. The control variables hold constant the initial economic conditions of local labor markets. These variables are meant to control for characteristics that may affect future changes in wage, employment, or sorting through domestic supply. The specification draws heavily from Dauth et al. (2014).

In turn I briefly discuss each control variable. The first control variable is the initial size of the local labor market in terms of employment. If there are different skill distributions on the worker side or productivity distributions on the firm side in small labor markets, then labor market size has the potential to affect sorting and possible the change in sorting over time. Labor market size

Table 1.9: All Coefficients for Main Specification of Equation (1.15)

	IV: LMR1 (1)	IV: LMR2 (2)
Change in export exposure	0.0105***	0.0080***
	(0.0026)	*(0.0025)*
Change in import exposure	-0.0039	-0.0014
	(0.0040)	*(0.0050)*
Initial local labor market employment	-5.75E-07**	-2.50E-07
	(2.67E-07)	*(3.89E-07)*
Initial local labor market sorting	-0.5506***	-0.6185***
	(0.0474)	*(0.0670)*
Initial share of employment in business service industries	-0.0005	-0.0005
	(0.0036)	*(0.0044)*
Initial share of business service occupations in business service industries	0.0016**	0.0018*
	(0.0008)	*(0.0009)*
Initial share of employment in non-auto manufacturing	-0.0004	-0.0001
	(0.0005)	*(0.0008)*
Initial share of employment in auto manufacturing	-0.0002	-0.0005
	(0.0007)	*(0.0008)*
Initial share of employment with a university degree	0.0059***	0.0054***
	(0.0017)	*(0.0019)*
Initial foreign-born share of employment	0.0018	-0.0002
	(0.0013)	*(0.0027)*
Initial female share of employment	0.0015***	0.0017***
	(0.0004)	*(0.0005)*
# geo fixed effects	74	214
Adj R2	0.371	0.445
N (county-periods)	650	650

Notes: All 2SLS regressions are weighted by the initial size of the regional labor force. Standard errors are clustered at the LMR2 level. Standard errors in parentheses. Main specification is column (2).

appears to have a negative effect on the future growth of sorting, but this coefficient is statistically insignificant from zero at the 10% level in the main specification. The second control variable is the initial strength of sorting with a local labor market. Given that my sorting measure is a correlation and, hence, bounded by definition, changes in sorting may be non-linear. For instance, if a region starts with a high level a sorting, it may be expected to experience less growth in sorting in the future. In fact, my results confirm this pattern with a strong negative coefficient on initial sorting.

Next I control for two measures that are relevant to outsourcing: the share of employment in business service industries and the share of business service occupations in business service industries. The second variable is a measure of outsourcing used in Goldschmidt and Schmieder (2017). Business service occupations such as logistics, cleaning, food, and security services are vulnerable to outsourcing arrangements such that a firm subcontracts these services to a business service firm. These employees often perform the same duties in the same location, but are employed by a different employer. Goldschmidt and Schmieder (2017) document that outsourced workers face a decline their wages and establishment fixed effect and, as a result, the rise in outsourcing can explain about 8% of the rise in sorting. Although, the share of employment in business service industries has an insignificant effect, the most direct measure of outsourcing–the share of business service occupations in business service firms–has a positive and significant effect on sorting. This result suggests that over this time period, regions with high initial levels of outsourcing were more likely to see an increase in their rate of outsourcing and, hence, increase sorting.

Next I control for measures of manufacturing concentration. As noted by Autor et al. (2013), it is important to control for the initial concentration of manufacturing employment so that comparisons of the effect of trade are made *within* the manufacturing sector. Since manufacturing employment is in long-term decline, it is important not to conflate the factors associated with the decline of manufacturing with changes in sorting patters. Given the prominence of the auto manufacturing sector in Germany, following Dauth et al. (2014) I break manufacturing employment into two components: non-auto and auto manufacturing. This helps to specifically address the concern that the results are driven by differential exposure to auto manufacturing trends. There are clear manufacturing trends in terms of employment and wages (employment is declining, wages are rising), but the prediction for sorting is less obvious. In general, manufacturing workers tend to be middle- to high-skilled and manufacturing firms tend to pay high firm premiums. Therefore, a reduction in manufacturing employment could result in less sorting as employment at the top end of the joint distribution is reduced. On the other hand, if the reduction in manufacturing employment is the result of improvements in technology that are complementary with skill, then sorting within the manufacturing industry may be increasing over time. The net result of these effects in an empirical question. As both manufacturing coefficients are insignificant, the results of Table 1.9 suggest that there is no trend in manufacturing sorting over time.

The final three control are standard and important measures of the composition of labor supply: education, nationality, and gender. The composition of labor supply can interact with firm types in complex ways to affect sorting. The results point to significant positive effects of the share of employment with a college degree and female. The interpretation of these results is that local labor markets with initially high level of college educated and female employment are more likely to see future increases in sorting.

Table 1.10 presents the same results as Table 1.4 except for females only. Over the period 1985 to 2009 sorting rose less dramatically for women compared to men. Table 1.1 shows that the total change in the correlation of worker and establishment fixed effects was 0.095 for women versus 0.231 for men. Given that sorting rose less for women, we may expect export exposure to have a weaker effect on female sorting. The main specification (reported in Column (6)) presents suggestive evidence in favor this view as the coefficient on export exposure is 0.0066 for women compared with 0.0080 for men. However, I cannot rejected the null hypothesis that these coefficients are the same. This result suggests that trade with Eastern Europe and China can account for a greater share of the rise in sorting for women than for men.

Table 1.10: 2SLS Results of Sorting on Trade Shocks for Females

	Region fixed effect					
	OLS: None (1)	IV: None (2)	IV: State (3)	IV: LMR1 (4)	IV: LMR2 (5)	IV: LMR2 (6)
Export exposure	0.0067**	0.0073*	0.0046	0.0071**	0.0063**	0.0066**
	(0.0032)	*(0.0044)*	*(0.0039)*	*(0.0032)*	*(0.0032)*	*(0.0026)*
Import exposure	0.0017	0.0014	0.0034	0.0037	0.0034	-0.0007
	(0.0030)	*(0.0041)*	*(0.0042)*	*(0.0037)*	*(0.0044)*	*(0.0037)*
Initial sorting						-0.9006***
						(0.0037)
Labor market controls	N	N	Y	Y	Y	Y
# geo fixed effects	0	0	11	74	214	214
Adj R^2	0.033	0.027	0.030	0.064	0.089	0.409
N (county-periods)	650	650	650	650	650	650

Notes: All 2SLS regressions are weighted by the initial size of the regional labor force. Standard errors are clustered at the LMR2 level. Labor market controls include: % employment in manufacturing, % high skilled, % foreign born, % female, and % routine occupation.

Table 1.11 presents the results of various robustness checks for the main results of the effect of trade exposure on local labor market sorting presented in Table 1.4. Columns (1) and (2) present the results for each interval separately. With only one observation per local labor market, however, I am constrained to use broad geographic regions (LMR1) for geographic fixed effects as some of the more narrow regions (LMR1) consist of only one local labor market or county (the German term is *kreis*). Therefore, the coefficients of columns (1) and (2) should be compared with column (1) of 1.9 (coefficient on export exposure of 0.0105) which uses fixed effects at the comparable geographic level. The results show that the coefficient on export exposure is significant in both intervals. There is suggestive evidence that the effect is stronger in the first interval, but I cannot reject the null hypothesis that the coefficients are equal.

Table 1.11: Robustness of the Effect of Export Exposure on Local Labor Market Sorting

	Specification:					
	First Interval '88–'99 (1)	Second Interval '93–'06 (2)	Control for Job Flows (3)	Constant WFE (4)	Net Exposure Total (5)	Net Exposure EE vs.CH (6)
Export exposure	0.0246**	0.0161***	0.0086***	0.0066***	-	-
	(0.0118)	*(0.0054)*	*(0.0025)*	*(0.0020)*		
Import exposure	-0.0183**	-0.0050	-0.0016	-0.0012	-	-
	(0.0084)	*(0.0038)*	*(0.0050)*	*(0.0043)*		
Net trade exposure	-	-	-	-	0.0047	
					(0.0037)	
Net trade exposure Eastern Europe	-	-	-	-	-	0.0090**
						(0.0041)
Net trade exposure China	-	-	-	-	-	*0.0024*
						(0.0059)
Change in job flows per firm	-	-	-0.00011	-	-	-
			(0.00018)			
Labor market controls	Y	Y	Y	Y	Y	Y
# geo fixed effects	74	74	214	214	214	214
Adj R^2	0.266	0.346	0.443	0.436	0.440	0.446
N (county-periods)	325	325	650	650	650	650

Notes: All 2SLS regressions are weighted by the initial size of the regional labor force. Standard errors are clustered at the LMR2 level. Standard errors in parentheses. Result are with respect to men only.

Column (3) presents the results of an important robustness check related to limited mobility bias. Table 1.3 along with previous results from Dauth et al. (2014) show that export exposure increases employment in local labor markets that experience export exposure shocks–particularly increases in manufacturing employment. As discussed in Section 1.3 limited mobility bias result in a downward bias in the correlation of worker and establishment fixed effect which depends on the number of job switchers per firm. As the number of job switchers increases the bias attenuates and, hence, the correlation increases. Hence, a concern for my estimation results is that the increased employment induced by export exposure leads to a greater number of job switchers per firm and hence an increase in the correlation caused by a reduction in the econometric bias rather than due to economic conditions.

Fortunately, the IAB grants access to job flows at the firm level based on the full universe of employment histories. This allows for an accurate count of the number of gross job changes per firm in each interval. To test the plausibility of this hypothesis I use the same research design in equation (1.15) except I insert the average number of gross job flows per firm as the dependent variable. The results show a significant coefficient on both export and import exposure at 5.340 (1.275) for export exposure and -2.236 (1.0116). Therefore, the concern that limited mobility bias may be driving the results is plausible.

In order to address this concern, I directly control for the change in job flows per firm in the regression of sorting on trade exposure based on equation (1.15). The result of this regression of presented in column (3) of Table 1.11. The coefficients on both export and import exposure essentially remain the same. Furthermore, the coefficient on the change in job flows per firm is statistically insignificant. The interpretation of this result is that local labor markets that saw a greater rise in job flows did not experience a rise in the correlation of worker and establishment fixed effects. This suggests that either the changes in average job flows per firm are too small to make a difference in terms of limited mobility bias or that the scope of limited mobility bias is small with these fixed effects. Given that the fixed effects are merged from Card et al. (2013) and based on the full universe, it is plausible that limited mobility bias is attenuated due to the large number of movers per firm that the full universe offers.

Column (4) address the question of whether the results are driven by changes in worker fixed effects across estimation intervals. In this specification, I create an alternative definition of the worker fixed effect. First, I de-mean all worker fixed effects within a given estimation interval. Next, for each worker I take the average worker fixed effect across all observation so that the worker fixed effect is constant over the life-cycle. I then use this alternative, constant worker fixed effect to define sorting in each local labor market. The results suggest that the coefficient of export exposure is slightly diminished with this alternative definition of worker fixed effects, suggesting that the changes in worker fixed effects across intervals are responsible for some of the total effect. However, I cannot reject the null hypothesis that the export coefficient is the same with the alternative definitions of worker fixed effects.

Columns (5) and (6) seek to understand the differential effects of trade exposure between Eastern Europe and China. Given that import exposure from Eastern Europe is highly correlated with import exposure from China (and also for export exposure), I cannot include components for both import and export exposure from both Eastern Europe and China in the same regression. Therefore, I use a measure of net exposure which is simply export exposure minus import exposure for Eastern Europe and China separately. Column (5) reports the results for net exposure including both Eastern Europe and China as a reference point. Column (6) presents the coefficients in a regression where both export to Eastern European and Chinese trade are included simultaneously. We can see that exposure from Eastern Europe has a stronger effect than exposure from China with a statistically significant coefficient of 0.0090 for Eastern Europe versus an insignificant 0.0024 for China. This result is consistent which evidence from Dauth et al. (2014) which shows that trade exposure from Eastern Europe has a stronger effect on employment and wages in Germany than trade exposure from China.

1.11 Appendix: A Model of Sorting within the AKM Framework

1.11.1 Model Setup

Suppose there is continuum of heterogeneous worker types x and firm types y where $x, y \in (0, 1)$. The output when a worker of type x and a firm of type y are matched is equal to $f(x, y) = xy$. Therefore, firm and worker types are complements as the production function is supermodular, i.e. ${\partial^2 f(x,y)}/{\partial x \partial y} > 0$. Labor is supplied inelastically with a total supply of m_x for each worker type.

Search frictions prevent firm and workers from instantaneously matching. A worker of type x receives a job offer from a firm of type y at a rate of λr_{xy} where λ represents a search friction and r_{xy} represents recruitment effort of firm type y for worker type x. Therefore, firms can choose to put more effort into finding workers of a particular type. Let $R_x = \sum_y r_{yx}$ denote the total amount of recruiting for workers of type x.

I assume a convex cost of recruiting and allow recruitment cost to be a function of both firm and worker type. Recruitment costs are given by $r_{xy}^2 c(x, y)$, where $c(x, y)$ represents a cost function. This cost function captures, in a reduced form fashion, the potential for differential recruitment costs by firm and worker type. For instance, this allows for the possibility that prestigious firms are more easily able to recruit highly productive workers. Two potential sources of lower recruitment costs for high-type firms seeking high-type workers are referral networks and preferences for amenities. If high-type firms initially hire high-type workers then it will be easier to locate other high-type workers if they share common social networks. If high-type workers have a greater preference for working at high-type firms due to, say, prestige or non-wage amenities then the reduced cost of recruitment can result from increased search effect on behalf of workers.

Once a worker and firm meet, they determine a wage to be paid to the worker, $w(x, y)$. For simplicity, I assume all workers come from unemployment and abstract from on-the-job search. Firms and workers engage in bargaining resulting in a wage setting rule of $w(x,y) = \beta xy$.[46] Jobs are destroyed at an exogenous rate δ. This simple wage equation maps into AKM as it is log separable in firm and worker wage components. Furthermore, the wage components map to productivity types.

Profit in period t is:

$$\pi_y^t = \sum_x \left[u_x^{t-1}\lambda r_{xy}^t + e_{xy}^{t-1}(1-\delta)\right] xy(1-\beta) - c(x,y)\left(r_{xy}^t\right)^2 \tag{1.18}$$

where u_x^{t-1} is the stock of unemployed of type x in period $t-1$ and e_{xy}^{t-1} is the number of workers of type x that firm y hired in period $t-1$. The costs of recruitment are convex in r_{xy} which ensures a unique solution.

1.11.2 Optimal Employment

The steady state aggregate flows out of unemployment must equal aggregate flows into unemployment:

$$\begin{aligned} \sum_y \lambda r_{xy} u_x &= \delta(m_x - u_x) \\ \Leftrightarrow \lambda R_x u_x &= \delta(m_x - u_x) \\ \Leftrightarrow u_x &= \frac{\delta}{\lambda R_x + \delta} m_x \end{aligned} \tag{1.19}$$

[46]This is equal to Nash bargaining where the outside option of each party is equal to zero. In a world without capacity constraints in terms of employment the cost of a vacancy will be zero, so this is reasonable on the firm side. On the worker side, this is equivalent to assuming that the worker prefers work at any wage over unemployment. The presence of a minimum wage or scarring effects from unemployment help to make this assumption more plausible.

The steady state flows into a firm must equal the steady state flows out of a firm:

$$\begin{aligned}
\lambda r_{xy} u_x &= \delta e_{xy} \\
\Leftrightarrow e_{xy} &= \frac{\lambda}{\delta} r_{xy} u_x \\
\Rightarrow e_{xy} &= \frac{\lambda}{\lambda R_x + \delta} r_{xy} m_x
\end{aligned} \tag{1.20}$$

Total output is the sum of each match output. Therefore, to maximize steady steady profits the firm chooses the optimal recruiting effort in each worker sub-market. I assume that each firm is relatively small and, therefore, takes the aggregate level of recruitment in the labor market R_x as given. The firm chooses recruitment effort to maximize profits for each worker type:

$$\max_{r_{xy}} \sum_x \frac{\lambda}{\lambda R_x + \delta} r_{xy} m_x yx \left(1 - \beta\right) - c\left(x, y\right) r_{xy}^2 \tag{1.21}$$

For each labor market there is a first order condition:

$$\begin{aligned}
\left[r_{xy}\right] \quad &: \quad \frac{\lambda}{\lambda R_x + \delta} m_x xy \left(1 - \beta\right) = 2c\left(x, y\right) r_{xy} \\
&\Leftrightarrow \quad r_{xy} = \frac{1}{2} \frac{\lambda}{\lambda R_x + \delta} m_x \frac{xy}{c\left(x, y\right)} \left(1 - \beta\right)
\end{aligned} \tag{1.22}$$

Inserting equation (1.22) in equation (1.20) yields optimal employment of each labor type at each firm of:

$$e_{xy} = \frac{1}{2} \left(\frac{\lambda m_x}{\lambda R_x + \delta} \right)^2 \frac{xy}{c\left(x, y\right)} \left(1 - \beta\right) \tag{1.23}$$

1.11.3 Sorting

Positive sorting or assortative matching results when high type firms hire a large share of high type workers as a portion of their total employment. A sufficient condition to ensure sorting that the derivative of the ratio of employment shares between a high and low productivity firm types increases as worker type increases. Employment share of worker type x in firm type y can be expressed as:

$$\begin{aligned}\frac{e_{xy}}{e_y} &= \frac{\frac{1}{2}\left(\frac{\lambda m_x}{\lambda R_x+\delta}\right)^2 \frac{xy}{c(x,y)}\left(1-\beta\right)}{\int \frac{1}{2}\left(\frac{\lambda m_z}{\lambda R_z+\delta}\right)^2 \frac{zy}{c(z,y)}\left(1-\beta\right) f_x(z)dz} \\ &= \frac{\left(\frac{\lambda m_x}{\lambda R_x+\delta}\right)^2 \frac{x}{c(x,y)}}{\int \left(\frac{\lambda m_z}{\lambda R_z+\delta}\right)^2 \frac{z}{c(z,y)} f_x(z)dz}\end{aligned} \tag{1.24}$$

where $f_x(z)$ is the density function over worker types. Let $y' > y$, the ratio of employment shares can be expressed as:

$$\begin{aligned}\frac{e_{xy'}}{e_{y'}}/\frac{e_{xy}}{e_y} = \frac{e_y}{e_{y'}}\frac{e_{xy'}}{e_{xy}} &= \frac{\int \left(\frac{\lambda m_z}{\lambda R_z+\delta}\right)^2 \frac{z}{c(z,y)} f_x(z)dz}{\int \left(\frac{\lambda m_z}{\lambda R_z+\delta}\right)^2 \frac{z}{c(z,y')} f_x(z)dz} \frac{\left(\frac{\lambda m_x}{\lambda R_x+\delta}\right)^2 \frac{x}{c(x,y')}}{\left(\frac{\lambda m_x}{\lambda_x R_x+\delta}\right)^2 \frac{x}{c(x,y)}} \\ &= \frac{\int \left(\frac{\lambda m_z}{\lambda R_z+\delta}\right)^2 \frac{z}{c(z,y)} f_x(z)dz}{\int \left(\frac{\lambda m_z}{\lambda R_z+\delta}\right)^2 \frac{z}{c(z,y')} f_x(z)dz} \frac{c\left(x,y\right)}{c\left(x,y'\right)}.\end{aligned} \tag{1.25}$$

A sufficient condition for sorting is that the ratio of shares is increasing in worker type. Taking the derivative with respect to x yields:

$$\frac{\partial^{e_y}/_{e_{y'}}{}^{e_{xy'}}/_{e_{xy}}}{\partial x} = \frac{\int \left(\frac{\lambda m_z}{\lambda R_z+\delta}\right)^2 \frac{z}{c(z,y)} f_x(z)dz}{\int \left(\frac{\lambda m_z}{\lambda R_z+\delta}\right)^2 \frac{z}{c(z,y')} f_x(z)dz} \frac{\partial \frac{c(x,y)}{c(x,y')}}{\partial x}. \tag{1.26}$$

Therefore, sorting results if the cost function satisfies the following condition:

$$\begin{aligned} & \frac{\partial \frac{c(x,y)}{c(x,y')}}{\partial x} > 0 \\ \Leftrightarrow \quad & \frac{c\left(x,y'\right)c_x\left(x,y\right) - c\left(x,y\right)c_x\left(x,y'\right)}{\left[c\left(x,y'\right)\right]^2} > 0 \\ \Leftrightarrow \quad & \frac{c\left(x,y'\right)c_x\left(x,y\right)}{c\left(x,y\right)c_x\left(x,y'\right)} > 1. \end{aligned} \tag{1.27}$$

For example, if $c(x, y) = \ln(xy)$ then

$$\frac{c\left(x,y'\right)c_x\left(x,y\right)}{c\left(x,y\right)c_x\left(x,y'\right)} = \frac{\ln\left(xy'\right)}{\ln\left(xy\right)} > 1 \quad \text{as} \quad y' > y. \tag{1.28}$$

The general condition is that the cost function must be log submodular. This ensures that high-productivity firms can recruit high-productivity workers at a relatively lower cost. If the cost function is the same as the production function, e.g. $c(x, y) = xy$, then this expression is equal to one and there is no sorting. In this case all firms will have the same worker-skill composition but more productive firms will be larger. If $c(x, y)$ is log supermodular, e.g. $c(x, y) = \exp(xy)$, then it is relatively more costly for high-productivity firms to recruit high-productivity workers and there will be negative assortative matching.

1.12 Appendix: Effect of Between- and Within-Group Changes in Employment on Changes in Sorting

An important consideration is whether changes in sorting are the result of changes in the relative size of employment shares across groups or due to the changing structure of employment within groups. By "groups" I have in mind industries, firm size groups, and worker flow groups. For example, since export exposure induces an relative increase in the size of the manufacturing sector, the increases in sorting may simply be the result of employment shifts to the manufacturing sector if manufacturing, in general, has a higher level of sorting.

To try to answer this question, I use the structure of my worker flow decomposition. However, in constructing counterfactual employment in each cell of the joint distribution I break the change in employment into two components. Let k denote a generic "group". Note that employment in the next period $(p+1)$ in each cell of the joint fixed effect distribution (ij) can be decomposed into what I will define as a between- and a within-group component. The between-component produces a counterfactual employment growth for each cell of the joint distribution equal the to the total employment growth for group k across all cell of the joint distribution multiplied by the initial employment share of cell ij from group k. This component capture the case in which employment growth occurs evenly across all cells according to the initial distribution so that the within-group distribution of employment is unchanged. The within-group component is then the remaining employment growth not explained by the between-group employment growth and represents changes in the structure of employment across the joint fixed effect distribution within a group. The following expression formalizes these components:

$$\begin{aligned}
\pi_{ij}^{p+1} &= \left[\pi_{ij}^{p} + \sum_{k} \frac{\triangle E_{ijk}}{E^{p}}\right] \frac{E^{p}}{E^{p+1}} \\
\pi_{ij}^{p+1} &= \left[\pi_{ij}^{p} + \sum_{k} \left(\frac{\triangle E_{ijk}}{E^{p}} - \frac{\pi_{ijk}^{p} \triangle E_{k}}{E^{p}}\right) + \sum_{k} \left(\frac{\pi_{ijk}^{p} \triangle E_{k}}{E^{p}}\right)\right] \frac{E^{p}}{E^{p+1}} \qquad (1.29)
\end{aligned}$$

where the first component represents the within-group component and the second, the between-group component.

To estimate the effect of between- versus within-group components, I follow a similar methodology to the worker flow decomposition as described in Section 1.4.1. To compute the between component I set the second component of equation 1.29 equal to zero for all cell of the joint distribution ij and groups k. The difference between change in the correlation with these counterfactual

Table 1.12: Decompostion of the Change in Sorting into Between- and Within-Group Components

	Change in correlation between worker and establishment fixed effects			
	I. Aggregate		II. Export-induced	
	BT-Group	WI-Group	BT-Group	WI-Group
Group definition	(1)	(2)	(3)	(4)
Industry	0.002	0.148	0.0000	0.0094
	(1.57)	*(98.43)*	*(0.06)*	*(99.94)*
Firm Size	0.004	0.144	0.0005	0.0089
	(2.89)	*(97.11)*	*(5.81)*	*(94.19)*
Worker Flow	-0.002	0.149	0.0003	0.0091
	(-1.12)	*(101.12)*	*(3.51)*	*(96.49)*
Industry*Firm Size	0.007	0.142	0.0004	0.0090
	(4.86)	*(95.14)*	*(4.58)*	*(95.42)*
Industry*Worker Flow	0.001	0.129	-0.0003	0.0074
	(0.51)	*(99.49)*	*(-4.87)*	*(104.87)*
Firm Size*Worker Flow	0.003	0.126	0.0004	0.0066
	(2.17)	*(97.83)*	*(6.04)*	*(93.96)*
Industry*Firm Size*Worker Flow	0.006	0.143	-0.0003	0.0098
	(4.28)	*(95.72)*	*(-3.38)*	*(103.38)*

Notes: The contribution of each component as a percentage of the total change is in parentheses. The total change in correlation for aggregate (export-induced) employment changes is 0.158 (0.0093). However, when performing the between-/within-group decomposition the components don't necessarily always add up to the total change in correlation as the decomposition is not strictly additively separable. Therefore, in the table, the total change for each row is defined as the within- plus between-group change. "Industry" consists of two groups: manufacturing and non-manufacturing. "Firm Size" consist of three groups: non-continuing firms, continuing firms below the median of the employment-weighted firm size distribution in the initial interval, and continuing firms above the median of the employment-weighted firm size distribution in the initial interval. "Worker Flow" consists of six groups: unemployment to employment, "other" to employment, labor market entry, job stayers, job-to-job between local labor market, and job-to-job within local labor market.

employment cells and the total change in the correlation serves as the estimation of the contribution of between-group changes in employment to the change in sorting.

Table 1.12 presents the results of this exercise. Panel I shows the results for the aggregate (or total) change in sorting and Panel II shows the results for the exogenous, export-induced portion of the change in sorting only. Results are reported for every possible combination of three group types: industries, firm sizes, and worker flows. Industries are defined broadly as manufacturing versus non-manufacturing. Given that I am using a 2% sample and I need every group to be populated in each local labor market to make consistent inferences, I am restricted to using this broad industry definition. Firm size consists of three groups: non-continuing firms, continuing firms below the median of the employment-weighted firm size distribution in the initial interval, and continuing firms above the median of the employment-weighted firm size distribution in the initial interval. Worker flows consists of six groups: unemployment to employment, "other" to employment, labor market entry, job stayers, job-to-job between local labor market, and job-to-job within local labor market.

The main result of Table 1.12 is that for all groups the vast majority of the change in correlation is due to within-group changes in the structure of employment across the joint firm and worker fixed effect distribution, rather than the changes in employment levels between groups. This is true for both aggregate employment changes as well is export-induced employment changes. This results suggests that a simple story such as changing employment shares across industries over time is not persuasive explanation for the rise of sorting.

1.13 Appendix: Effect of Sequencing on Worker Flow Decomposition Results

As in other similar methods, the decomposition method laid forth in 1.4.1 is potentially sensitive the order in which employment changes of different groups are varied. For example suppose a generic function f depends on e_k^p which represents employment in worker flow categories k at time p. Given six worker flow categories the following would be equally valid represents of the contribution of flow 1 to the total change in f:

$$\hat{\alpha}_1^1 = f\left(e_1^{p+1}, e_2^p, e_3^p, e_4^p, e_5^p, e_6^p\right) - f\left(e_1^p, e_2^p, e_3^p, e_4^p, e_5^p, e_6^p\right)$$

or

$$\hat{\alpha}_1^2 = f\left(e_1^{p+1}, e_2^{p+1}, e_3^p, e_4^p, e_5^p, e_6^p\right) - f\left(e_1^p, e_2^{p+1}, e_3^p, e_4^p, e_5^p, e_6^p\right).$$

In each case only flow one is varying while the other flows are held constant. If the arguments of f are not additively separable, however, there is no guarantee that $\hat{\alpha}_1^1 = \hat{\alpha}_1^2$. In the case that f is a correlation across a joint distribution of employment, the arguments are not additively separable. Therefore, the sequence in which I compute the contribution of worker flows to sorting may matter.

Table 1.13 reports summary statistics of the results of the worker flow decomposition over all possible sequences for the aggregate changes in sorting and the export-induced changes in sorting. The results can be comparable to results of the main specification reported in Table 1.5. In this case summary statistics are computed over 32 different sequences. This exercise suggests that the main results are not sensitive to changes in sequencing. Even if we take the share of sorting for labor market entry at the minimum level across all sequences it is still greater than the maximum of any other share. Therefore, the importance of labor market entry is robust to sequencing. This is particularly true for the case of export-induced changes in sorting which shows very little variance with respect to the sequencing of the decomposition.

Table 1.13: Descriptive Statistics of Worker Flow Decomposition Across Different Sequences

	I. Aggregate				II. Export-Induced			
	Mean (1)	S.E. (2)	Max (3)	Min (4)	Mean (5)	S.E. (6)	Max (7)	Min (8)
Unemployment to Employment	0.0031	0.00024	0.0052	0.0012	0.00003	0.000001	0.00004	0.00003
	(1.93)	*(0.15)*	*(3.26)*	*(0.77)*	*(0.35)*	*(0.01)*	*(0.41)*	*(0.28)*
"Other" to Employment	0.0152	0.00025	0.0180	0.0130	0.00093	0.000001	0.00094	0.00092
	(9.61)	*(0.16)*	*(11.38)*	*(8.20)*	*(9.92)*	*(0.01)*	*(10.02)*	*(9.83)*
Labor Market Entry	0.0883	0.00090	0.0980	0.0782	0.00448	0.000002	0.00450	0.00445
	(55.71)	*(0.57)*	*(61.85)*	*(49.38)*	*(47.93)*	*(0.02)*	*(48.19)*	*(47.68)*
Job Stayers	0.0203	0.00076	0.0271	0.0143	0.00234	0.000001	0.00235	0.00233
	(12.79)	*(0.48)*	*(17.11)*	*(9.02)*	*(25.08)*	*(0.01)*	*(25.19)*	*(24.97)*
Job-to-Job Between Region	0.0216	0.00036	0.0256	0.0183	0.00156	0.000002	0.00158	0.00155
	(13.63)	*(0.23)*	*(16.15)*	*(11.56)*	*(16.73)*	*(0.02)*	*(16.92)*	*(16.55)*
Job-to-Job Within Region	0.0101	0.00034	0.0135	0.0072	-0.00001	0.000001	0.00000	-0.00001
	(6.35)	*(0.21)*	*(8.49)*	*(4.52)*	*(-0.06)*	*(0.01)*	*(-0.01)*	*(-0.11)*

Notes: The contribution of each component as a percentage of the total change is in parentheses. The total change in correlation for aggregate (export-induced) employment changes is 0.158 (0.0093). There are 32 different sequences by which the six worker flows can be ordered to compute counterfactual employment distributions. "S.E." denotes the standard error across the 32 sequences.

Chapter 2

The Effects of Employment Shocks on the Self-Employed: Evidence from Natural Experiments in Professional Golf

I thank the PGA TOUR for providing access to ShotLink™data..

2.1 Introduction

Temporary employment shocks can have significant and lasting consequences for workers' future earnings prospects. For example, individuals who graduate college during a recession or are displaced from a job suffer large initial earnings losses than can persist for more than a decade.[1] Yet,

[1] Persistent earnings losses from job displacement are robustly documented in, among others, Topel (1990), Ruhm (1991), Farber (1993, 1996, 2015, 2017), Stevens (1997), von Wachter et al. (2009), Couch and Placzek (2010), Davis and von Wachter (2011), and Lachowska et al. (2018). For surveys of the job displacement literature see Hamermesh (1989), Fallick (1996), Kletzer (1998), von Wachter (2009), and Carrington and Fallick (2017). As an example of an estimate of the magnitude of losses with a representative sample of US workers, von Wachter et al. (2009) find initial losses due to job displacement reaching 30% of earnings, with losses of around 20%, 20 years later.

despite a wide variety of studies for different populations and countries, the literature focuses almost entirely on the wage-and-salary workers. As a result, it is unclear whether self-employed workers face similar consequences from temporary employment shocks. But, it is important to understand earnings dynamics for this group as the self-employed account for about one tenth of the workforce and are integral to business creation (Parker, 2004; Hipple, 2010). Furthermore, learning about the response of the self-employed to employment shocks provides a new perspective from which to understand persistent earnings losses of wage-and-salary workers.

Theories of how temporary employment shocks cause persistent earnings losses focus either on the role of human capital depreciation or the role of search and contracting frictions between firms and workers. On the one hand, human capital theory suggests that after an employment shock, as workers are either unemployed or missallocated to poor jobs, they are susceptible to human capital depreciation and a reduction in productivity growth potential. On the other hand, substantial dispersion in firm-specific wage policies may result in some firms offering jobs that pay higher wages. If employment shocks displace workers from high paying firms, strong job search frictions may impede earnings recovery and produce persistent earnings losses without changes to workers' underlying productivity. For the self-employed, the frictions and constraints induced by the joint decision making process between firms and workers are not operative as these workers are not hired by firms. However, disruptions to the human capital or skill accumulation process may have similarly adverse consequences for future productivity and earnings. Therefore, if frictions between workers and firms cause persistent losses then we should expect to see a relatively quick recovery for self-employed workers facing temporary employment shocks. However, if productivity depreciation causes erosions in earnings, then we still could see persistent losses for the

---

Multiple studies have also found significant earnings losses for individuals who graduate college during recessions (Kahn, 2010; Oreopoulos et al., 2012; Oyer, 2006, 2008; Genda et al., 2010). For example, for a 5% increase in the unemployment rate (a typical recession), Kahn (2010) finds a 30-35% wage loss at labor market entry which attenuates only moderately to 12.5% after 15 years. Oreopoulos et al. (2012) find less severe wage losses in Canada of nine percent, initially upon labor market entry, with losses lasting for around ten years.

self-employed.

To enhance our understanding of the consequences of temporary employment shocks, I analyze a labor market of self-employed professional golfers. This setting offers unique features which facilitate clean identification of the effects of employment shocks on both future earnings and productivity. First, earnings and employment outcomes are available for each golfer over a long duration. This data allows for the construction of a panel data set capable of tracking golfers' outcomes for up to sixteen years past an employment shock event. Second, entry rules set by the *PGA TOUR*, the top professional golf league, create some compelling regression discontinuity designs in which annual employment rights are effectively randomized across golfers near the treatment threshold. These large, temporary, and exogenous employment shocks create excellent variation with which to identify the causal effect of employment shocks on long-term outcomes. Third, administrative golf data includes directly observable measures of productivity. Previous research had has difficulty distinguishing between the mechanisms which produce persistent earnings effects partly because measures of productivity are rarely available. In contrast, in the golf setting I am able to use the regression discontinuity designs to directly estimate the effect of an employment shock on productivity. Therefore, we can at the same time be confident that selection issues do not bias the estimates and use the results on golfer productivity to better understand the potential mechanisms underlying the earnings response.

I use two separate natural experiments to estimate the causal effect of temporary employment shocks and future earnings, employment, and productivity. These natural experiments are byproducts of the PGA TOUR's entry rules. To maintain a highly competitive level of play, each year the PGA TOUR rewards the best performing golfers on a lower golf tour with annual employment rights. At the same time, the worst performing PGA TOUR members have their employment rights revoked. These entry rules are implemented in such a way to produce sharp discontinues in membership benefits around a threshold. For instance, prior to 2012 there were two primary ways for golfers to advance to the PGA TOUR. First, a golfer could place among the top positions on the season ending money list of the developmental tour, the *Web.com Tour Money List (ML)*. Second, a player could place among the top positions in an extended qualifying tournament known as the PGA TOUR Qualifying Tournament or *Q School*. In both cases, entry rules create sharp discontinuities in membership benefits at a discrete threshold such that those who place just below the threshold enjoy PGA TOUR membership for the next year, while those who place just below the threshold do not. Therefore, these discrete treatment thresholds provide excellent conditions to utilize a regression discontinuity design to identify the effects of employment shocks on future outcomes.

My results can be summarize in three key findings. First, I estimate very large effects of treatment, an annual PGA TOUR membership, on initial earnings and employment–suggesting that these experiments do, in fact, produce large employment shocks. Narrowly crossing the treatment threshold is estimated to provide a golfer with a 156% and 90% earnings increase and access to 17 and 21 additional PGA TOUR events in the subsequent year for the Web.com Tour ML and Q School natural experiment, respectively. Second, despite these large initial earnings differences, golfers are able to quickly recover. After just three years, earnings of the treatment and control groups equalize. Furthermore, earnings differences between the groups remain statistically indistinguishable from zero for the remainder of the sixteen-year observation period. Third, I find no productivity response to the employment shocks. After adjusting for competitor quality, there is no difference in the performance between treatment and control groups both in the initial years after the employment shock or in subsequent years. This result also holds for younger golfers–a group that may be particular vulnerable to disruptions in the skill accumulation process.

In general the results of the two experiments, Q School and Web.com Tour ML, are very similar. This is in spite of some material differences in the characteristics of the potential treatment populations. For instance, the Web.com Tour ML sample has higher average ability and less dispersion in ability than the Q School sample. The robustness of the results across the two treatments lends credibly to the estimates and suggests that the effects may be similar across a wider range of ability levels than the specific treatment populations addressed by the natural experiments.

Although the PGA TOUR employment shocks create large initial earnings losses for the control group, these losses quickly dissipate. This result suggests that professional golfers, and possibly the self-employed more generally, are able to recover more quickly from employment shocks than wage-and-salary workers. But, these results beg the question as to why golfers' earnings recover so quickly despite facing such a large initial shock. As alluded to previously, the theories for why wage-and-salary workers face persistent losses from employment shocks can be classified into explanations based on frictions or constraints between firms and workers and explanations based on human capital or skill depreciation directly leading to lower productivity. A consistent explanation for both sets of results is that firm rent-sharing policies combine with firm-worker frictions to produce persistent earnings losses from employment shocks. As a corollary, the degree of persistence will depend on the strength of the frictions. In fact, the compensation policies of golf tours appear to effectively produce rent-sharing policies. Evidence for this interpretation can be found in the result that, in the first year after treatment, both treatment and control groups perform very similarly, yet the treated group enjoys a large pay raise. Thus the PGA TOUR offers significant "rents" in the sense that compensation is greater for the same level of performance, at least for golfers of marginal ability. My preferred interpretation of the results is that, although there is potential for rent-sharing in professional golf, frictions are minimized and, hence, the effect of employment shocks on future earnings is much less persistent.

In an effort to substantiate the claim that frictions are reduced in the professional golf labor market, I perform two exercises. First, using techniques from the earnings dynamics literature, I decompose earnings shocks into permanent and transitory components of total variation. I find a larger proportion of transitory variance than has been estimated with PSID and IRS earnings data for wage-and-salary workers. The larger transitory component suggests that temporary shocks do not persist and corroborate the interpretation of professional golf as a low friction environment. Second, I analyze transition rates on and off the PGA TOUR and compare them to transitions rates across firm types for wage-and-salary workers. I find high exit rates for the treated sample of golfers, i.e. golfers that are awarded a PGA TOUR membership are very likely to get demoted back to a lower tour in the future. In fact, these exit rates are much higher than the estimated probability of moving down the job ladder for wage-and-salary workers in the broader economy. This result provides an indication of a potentially important channel by which frictions may be different for wage-and-salary workers relative to the self-employed. In particular, it appears that the persistence of employment relationships, rather than job search frictions, may be an important mechanism for the persistent earnings losses of wage-and-salary workers.

The main contribution of this paper is to extend our understanding of the effects of temporary employment shocks, including the large literature on the effects of job displacement and entering the labor market during a recession. To my knowledge, this is the first study to focus on the effect of employment shocks on a population of self-employed workers. My findings suggest that self-employed populations may be able to recover from employment shocks more quickly than wage-and-salary workers. A key reason for this quick recovery is that golfers' performance does not suffer from a substantial employment shock. Given that researchers have found significant, negative health effects from job displacements (e.g. Sullivan and von Wachter, 2009), this is an interesting result as, ex-ante, we could have expected a decline in productivity due to either the psychological toll of narrowly missing a great opportunity or the reduction in peer quality associated with playing on a less-competitive golf tour. Furthermore, I add to this literature by distinguishing between two potential channels that impede earnings convergence after an employment shock. Both a low probability of finding a job good after a negative shock (job search friction) and a high probability of staying in a job after a positive shock (job security) can contribute to persistence earnings losses of negatively shocked workers. Relative to golfers, I find evidence that the job security of nondisplaced, wage-and-salary workers is a more important source of persistence than job search frictions facing the displaced. My findings are consistent with two recent papers in the macroeconomics literature which develop search models to decompose the sources of job displacement losses. Both Jarosch (2015) and Jung and Kuhn (2018) highlight the relativity greater job security of nondisplaced workers as an important factor in accounting for persistent earnings losses from job displacement.

I follow an extensive literature in labor economics which studies specific labor markets to better understand more general phenomena. As a particularly influential example, Baker, Gibbs and Holmstrom (1994a,b) investigate theories of internal labor markets and wage contracts by analyzing the wage policies of a single large firm. Golf, in particular, has been a fruitful area to apply economic theory due, in part, to the quality of PGA TOUR data. Prominent examples of golf studies include Guryan et al. (2009) who test for peer effects in the workplace using information on playing partners, Pope and Schweitzer (2011) who test for loss aversion with highly detailed scoring data, and Brown (2011) who studies the effects of superstars on competitors' performance. Other researchers haves studied specific labor markets to understand similar issues related to the role of luck in labor market outcomes. Oyer (2006, 2008) studies the impact of initial labor market conditions on long-term outcomes for economists and the impact of stock market fluctuations on career choice for MBA students. Bertrand and Mullainathan (2001) study whether CEO's are rewarded for random fluctuations in stock prices. Contrary to my results for professional golfers, these studies generally find an important and persistent role for luck in labor market outcomes. The disparity in these results is consistent with an interpretation of professional golf as a highly competitive environment in which the effects of transitory shocks do not persist.

This paper also adds to a literature studying the unique features of self-employment both by documenting the earnings response of self-employed professional golfers to employment shocks and by contrasting that response to the typical response documented for wage-and-salary workers. The self-employed account for about 10% of the workforce and have greater dispersion in income than wage-and-salary workers (Parker et al., 2005). In fact, self-employed workers are over-represented in both the upper and lower tails of the overall income distribution (Parker, 1997). Indeed the low-skill self-employed tend to have lower wages than low-skill wage-and-salary workers, while the high-skill self-employed tend to have higher wages than high-skill wage-and-salary workers (Krashinsky, 2008). In this paper, I focus on a population, professional golfers, located at the upper end of the income distribution. Many other studies use a comparison between self-employed and wage-and-salary workers to understand features of the labor market. For example, Krashinsky (2008) uses differential changes in the wage distribution to investigate the role of institutional factors in rising wage inequality, Lazear and Moore (1984) use differential experience gradients in wages to argue that incentive pay is used to overcome agency problems, Carrington et al. (1996) use differential business cycle variation in wages to understand the role of sticky wages in business cycle wage fluctuations, and Moore (1983) uses wage differentials to investigate employer discrimination.[2] A common thread throughout these studies is to cite the lack of firm-worker relationships as justification to interpret the self-employed labor market as an approximation of a competitive spot market. To my knowledge, no prior research has analyzed the effect of employment shocks on earnings dynamics of self-employed workers or the transitory and permanent components of self-employed earnings variation.

In what follows, Section 2.2 provides a brief background of theories of persistence losses from temporary employment shocks and how these theories apply in a golf setting. Section 2.3 provides a brief overview of relevant facts about golf as well as describes the data. Section 2.4 describes the natural experiments in detail. Section 2.5 states the RD identification assumptions, describes

[2]Beyond labor economics, Gruber and Poterba (1994) use changes in tax policy for the self-employed to estimate the price elasticity of demand for health insurance.

the empirical estimation, and defends the validity of the identification assumptions in this setting. Section 2.6 present the main results on earnings and productivity and then employment. Section 2.7 documents some key differences between professional golf and the broader US labor market. Finally, Section 2.8 concludes and discusses the implications of the findings.

2.2 Theoretical Background and Framework

2.2.1 Theories of Career Earnings Progression

In order to understand theories of the persistent effects of employment shocks, it is helpful to take a step back and understand theories of post-education earnings growth. The leading theories can generally be clustered into the three classes of general human capital accumulation, specific human capital accumulation, and job search. I briefly describe each below.

The seminal work of Becker (1964) and Mincer (1974) provides the foundation for much of human capital theory and also the distinction between general and specific human capital accumulation. In brief, general human capital is knowledge or skills that are transferable to many occupations, industries, tasks, or firms. In a slight abuse of terminology, consider three formulations of the general human capital accumulation process. First, and most simply, individual may inherit skills simply through the biological process of aging or maturing. Second, human capital may be accumulated with experience on the job without an explicit cost. This is referred to as *learning by doing* and was originally discussed by Arrow (1962) and Rosen (1972a). A final form of general human capital accumulation is on-the-job training where there is an explicit cost to human capital accumulation which normally manifests itself as time spent on non-production activity. The consequences of on-the-job training investment decisions are formulated in Becker (1964) and Ben-Porath (1967).

The early work on human capital by Becker also laid forth a view of human capital as specific to the employer. Firm-specific human capital creates theoretical difficulties in terms of who will pay for training given that there is no market for firm-specific skills. However, economists have developed this idea in many plausible forms such as industry-specific human capital (Neal, 1995), occupation-specific human capital (Kambourov and Manovskii, 2009), and task-specific human capital (Gibbons and Waldman, 2006). A clear implication of specific human capital accumulation is that if workers are forced to change jobs, they may lose the return to their specific skills.

A competing view of career progression is search theory. Given that it takes time to find a good job or a good match, we may see earnings growth due to movements up a job ladder which don't result from human capital accumulation. In fact, Topel and Ward (1992) find that one third of earnings growth for young men between the ages of 18 and 34 can be attributed to job changes. Whether these gains are the result of movements to higher paying firms in general or more productively matched worker-firm combinations is unclear. Human capital theory and search theory are not mutually exclusive. A theory which blends both is that different firms offer different human capital accumulation opportunities (Rosen, 1972b) and, hence, workers search to find the best learning environments.

2.2.2 Theories of Persistent Losses from Employment Shocks

With a brief background on the theories of career progression, let us consider why temporary employment shocks may cause persistent earnings losses. For purposes of exposition, I draw a distinction between the theories in terms of their implications for worker productivity. Both theories of general and specific human capital affect productivity, whereas theories of rent-sharing or worker-firm complementaries usually take worker productivity as constant.

As a first step, note that human capital theory, alone, is incapable of explaining persistent earnings effects. Consider the case of job displacement. On one hand, in a frictionless world there is no basis for displacement. Rather wages will instantly adjust to marginal productivity. But even given some random displacement shocks, workers should be instantly re-hired at their previous wage. Therefore, some search or contracting frictions are needed to supplement human capital theory in order to create persistence.

General human capital depreciation during unemployment can theoretically explain persistent earnings effects. For instance search frictions preventing quick re-entry into the labor market could cause significant depreciation. However, the job displacement literature finds sharp losses even for those quickly re-hired, raising doubts over the plausibility of this mechanism (von Wachter et al., 2009). A variant of the general human capital story is that the workplace environments of firms may vary in terms of the general human capital accumulation opportunities offered (Rosen, 1972b). Coupled with search frictions, displacement could result in a stagnation of earnings growth until the worker is able to get back to a learning environment of equal quality. In both cases, depreciation or stagnation, earnings losses accrue from productivity losses and the persistence of earnings losses depends on the strength of the frictions.

Specific human capital accumulation and search frictions can also theoretically explain persistent earnings losses. For instance, if workers accumulate occupation-, industry-, or task-specific skills with tenure, then any difficulty in finding a similar job after displacement will result in earnings losses. Once again, losses may persist depending on the frictions preventing re-hiring within the same occupation, industry, or task. If jobs disappear forever, then losses could be permanent.

Theories of rent-sharing can also explain persistent earnings losses. For instance, if there are good and bad jobs in the economy, it takes time to find the good ones, and displacement knocks workers down to the bottom of the job ladder, then persistent earnings losses will accrue. There are a few plausible options regarding what constitutes a "good job". For instance, firms that have market power or are highly productive or profitable may share some of these rents with their workforce. Alternatively, firms may take advantage of monopsony power to pay workers less than marginal productivity (Burdett and Mortensen, 1998). On the other hand, a "good job" could be a good match in which the type of skills the worker holds are well suited for the quality of the firm. In all cases, underlying worker productivity is not changing, but rather, disruptions while climbing the job ladder create earnings losses.[3]

Whether due to human capital, rent-sharing, or firm-worker matches, the persistence of earnings losses depends on the strength of the frictions. In fact all of the above theories share a similar structure. There is a good and a bad state of the world and an employment shock knocks a worker from the good to the bad state. The effects of a displacement, therefore, depend on the exit rate from and the entry rate into the good state. In general, the lower these transition rates, the more persistence the effects of an employment shock will be. The bad states in each of these theories correspond to either unemployment, a poor learning environment at work, a mismatched job (in terms of either industry, occupation, task, worker-firm production complementaries, etc.), or a generally low-paying firm.

A previously noted, a different type of explanation for displacement losses is that they result from biases in their estimation. An important challenge for studies of employment shocks is negative selection or the inability to create satisfactory counterfactual groups with which to compare earnings trajectories. For example, in the case of job displacement, those displaced may be lower quality workers along unobservable dimensions, workers may sort into firms according to the likelihood of a layoff, or high quality workers may leave firms prior to the mass displacement event. Another measurement issue concerns the nature of contracts. For instance, suppose that contracts are back-loaded so that pay rises over time without relation to a worker's productivity. These contracts may incentivize investments in firm-specific human capital (Salop and Salop, 1976) or disincentivize shirking (Lazear, 1981). In the case of job displacement, the back-loading of con-

[3]It may seem strange to insist that worker productivity does not change in a model with worker-firm production complementaries. After all, in this case, the marginal productivity of the worker depends on the firm. However, I wish to make the subtle distinction between specific human capital models where skills accrue with tenure versus worker-firm productive complementarity models where the underlying ability of workers is constant, but some workers are naturally better suited to some firms.

tracts creates a measured displacement effect only because wage contracts are being compared at different stages of job tenure.

Other explanations for persistent effects of temporary employments shocks which are less relevant to the golf setting are adverse selection in the labor market (Greenwald, 1986; Gibbons and Katz, 1991), entry conditions and bargaining power (Beaudry and DiNardo, 1991), and downwardly rigid wages (MacLeod and Malcomson, 1993).

2.2.3 Applicable Theories to Professional Golf

Analyzing employment shocks in the golf labor market is similar to studying the consequences of job displacement, but in a more straightforward environment. There are at least four important ways in which golf differs from a general labor market. First, the nature of production is simple and transparent. Output is objective and produced at the individual level. Production functions do not depend on the interactions between capital, managers, and coworkers. Second, the market highly competitive. Success in golf is very lucrative and, as a result, individuals compete worldwide to join the PGA TOUR. Yet, despite the large supply, the PGA TOUR places a limit on the number of golfers that can attain membership–further adding to the competitiveness of the environment. Third, there are no long term contracts or employee protections from firing. Prize money is awarded based on tournament performance and membership status is generally granted annually, although longer term exemptions are given for tournament victories. Fourth, there are no search frictions in the sense that there are only a handful of professional golf tours which are well known. In economic terms, therefore, we can think of golf as a highly competitive labor market with little regulation and a production function comprised of a single input, labor. All types of labor are perfect substitutes for one another, resulting in "firms", or golf tours, that are simply aggregations of independent production systems.

The two natural experiments resemble job displacement in the sense that the treated group enjoys a "good" job for at least a year and the control group is relegated to a "bad" job. In contrast to job displacement, however, unlucky golfers do not face unemployment risk since those that narrowly miss the cutoff are guaranteed membership on the Web.com Tour. Nevertheless, the spirit is similar in the sense that there is an exogeneous shock to the employment state. Since golf is a simplified labor market many of the variants of human capital and rent-sharing theory do not apply. However, I argue that there are plausible variants of both theories that apply to golf. Below I quickly describe which theories are not relevant and then I lay out the relevant mechanisms.

First, the controlled environment reduces measurement problems. The PGA TOUR entry rules create compelling natural experiments. In Section 2.5, I provide evidence that the identification assumptions for a RD design are satisfied. A strength of this setting is that it alleviates concerns over negative selection. Also, the earnings process is transparent so there is no uncertainty over the length of the employment contract. Hence, there is no concern over back-loading of contracts.

Second, there is limited role for specific human capital or tour-golfer complementarities. Although golfers, by definition, work in the same industry and occupation and perform the same tasks, it is possible that some golfers are better suited for the developmental tour, the Web.com Tour, whereas others have skills that are more conducive to the PGA TOUR. However, the data suggests that specific skills or production complementarities are unlikely. For instance, the correlation between Web.com Tour ML ranking in year t and PGA TOUR ML ranking in year $t+1$, for those who qualify from the Web.com Tour to the PGA TOUR, is 0.29. Whereas the correlation between PGA TOUR money list ranking in year t and $t+1$, for PGA TOUR played ranked between 101 and 125 in year t, is 0.18.[4] These correlations show that past success on the Web.com Tour is similarly predictive of future success on the PGA TOUR as past success on the PGA TOUR.[5]

[4] I try to select a comparable group of golfers to the Web.com Tour ML qualifiers and a comparable number of goflers. The top 125 on the PGA TOUR money list earn a tour card for the next year. Therefore, players ranked 101 through 125 are the 25 closest players to not being fully exempt on the PGA TOUR.

[5] The fact that these correlations are quite low suggests a large role for yearly variance in golf earnings.

Therefore, it does not appear that PGA TOUR members are more suited for PGA TOUR play than Web.com Tour members.

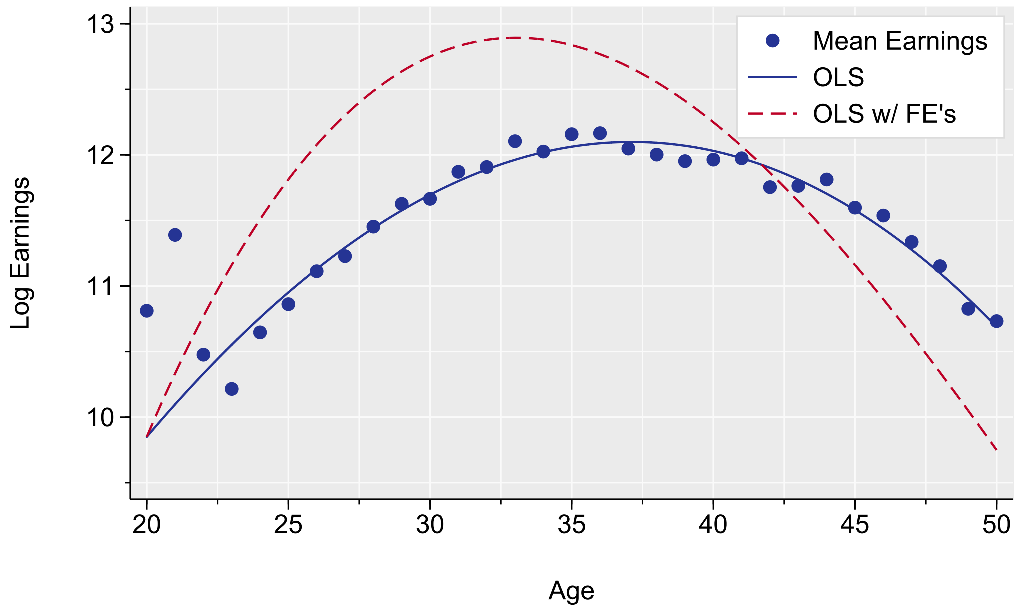
General human capital deterioration from unemployment does not apply as all individuals are guaranteed employment in the next year. One plausible hypothesis, however, is that treated golfers accumulate more general human capital or skill depending on their environment. In particular, by playing a full year on the PGA TOUR instead of the Web.com Tour, golfers may improve their performance. PGA TOUR members play with the best golfers in the world, for very high stakes, and on the most difficult golf courses in the world. Any one of these factors could plausibly contribute to a golfer's development. For instance, young golfers may learn from more skilled or experienced golfers. Also, high stakes, in terms of money or prestige, could incentivize golfers to invest more in their game. Finally, difficult golf courses and high pressure situations may give golfers valuable experience that can be applied in the future. The persistence of these potential effects will depend on both the differential amount of skill can that be learned on the PGA TOUR versus in the Web.com Tour and the exit and entry rates onto and off of the PGA TOUR.

Empirically there is a relationship between age, earnings, and performance. Figure 2.1 shows average earnings and scoring average over the life-cycle for golfers on the PGA TOUR and Web.com Tour. Each panel shows raw means (blue dots), an OLS regression fit of a cubic in age (solid blue line), and an OLS with individual fixed effects (dashed red line). Figure 2.1a and Figure 2.1b show substantial earnings and performance growth for young golfers. Therefore, sustained shocks to the determinants of performance over this period of performance growth may cause persistent earnings losses.

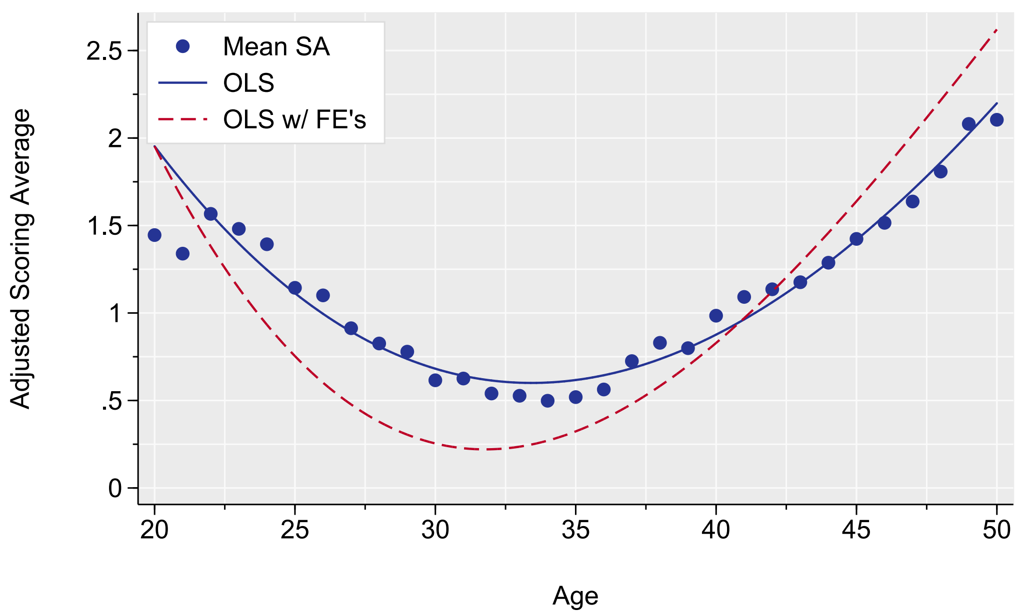
Another plausible hypothesis is a rent-sharing type of mechanism. If, conditional on the level of performance, the PGA TOUR awards more prize money relative to other tours, then the PGA TOUR can be said to share rents with its members. Potentially, this is similar to a rent-sharing relationship between a firm and its workers. Rent-sharing is consistent with recent evidence showing

Figure 2.1: Career Profiles in Professional Golf

(a) Earnings Profile: Log US Earnings



(b) Productivity Profile: US Adjusted Scoring Average



Notes: The sample includes all golfers with some positive earnings on either the PGA TOUR or the Web.com Tour from 1990 to 2014. The blue dots represent mean log earnings by age. The blue, solid lines represent the prediction line of an OLS regression of log earnings on a cubic in age. The red, dashed lines represent the prediction line of an OLS regression of log earnings on a cubic in age and individual fixed effects. The OLS with fixed effects is normalized to begin at the same level at the OLS without fixed effects.

that a large share of wage variation can be attributed to the identity of the firm (Card et al., 2013; Song et al., 2018). Although there are many differences between the firm-worker and tour-golfer relationship, they may be similar in terms of rent-sharing.[6] In contrast to a general human captial story, persistence under rent-sharing depends solely on the exit and entry rates from onto and off of the PGA TOUR.

The two theories, general human capital and rent-sharing, can be distinguished in terms of their predictions for future earnings and productivity. Under the human capital hypothesis, I expect golfers to learn something during their year on the PGA TOUR. Therefore, in future years the treated should outperform the non-treated. Since under the rent-sharing hypotheses there are no performance gains from playing on the PGA TOUR, I expect less persistent earnings effects and no effects on productivity. If the exit and entry transition rates from the PGA TOUR to the Web.com Tour are high, then I expect no persistence under the rent-sharing hypothesis.

Another dimension along which the theories diverge is in terms of their effects over the life-cycle. Similar to a general labor market, Figure 2.1a shows that earnings grow until about age 33 (OLS with fixed effects). However, instead of flattening out like typical labor market earnings, golfers' earnings fall quickly past their peak. The decline in earnings is almost as steep as the rise with 20 and 50 years olds earning about the same amount. This feature is likely to be the result of physical deterioration, a factor of less importance in the modern workforce. Figure 2.1b shows that similar patterns apply to performance. Focusing on the fixed effects specification, performance improves steadily when young, peaks around age 32, and then quickly declines. The productivity profile is similar to the earnings profile with the slight distinction that performance falls off more

[6]The golf tour-golfer relationship is unlike the firm-worker relationship in many economically significant ways. For instance, there is less scope for search frictions since knowledge of the tours is common, public, and well-advertised around the world. Also, tours neither make hiring decisions nor sign long-term contracts. Instead, they use objective entry rules in an effort to accept only the best players.

than earnings.[7] These figures suggest that young players learn with experience, but then at some point they physically depreciate. If golfers accumulate most of their skills early in their career, then the learning environment should be more important for younger golfers. Therefore, under the human capital hypothesis I expect to see greater and more persistent treatment effects for younger golfers.

2.3 Golf Background and Data

2.3.1 Relevant Facts about Professional Golf

The object of golf is to strike a ball with a club as few times as possible until it enters a hole. Score is kept over a course consisting of eighteen holes constituting a *round*. A golf tournament usually consists of four rounds played over four days. In a tournament there are typically 130 to 170 competitors. Usually the highest scoring half of the field is *cut*, or disqualified, after the first two rounds of play and do not earn any prize money.

Much like other sports, golf is played professionally for high stakes by individuals from around the world. The PGA TOUR is an organization which hosts some the world's most prestigious and lucrative golf tournaments. In 2012, the PGA TOUR held 45 official tournaments. The median golfer competed in 23 events and earned $618,628.50. Rory McIlroy was the leading money winner, earning $8,047,952, while the bottom 10% earned less than $33,960. J.J. Killeen played in the most tournaments, 33, while the bottom 10% played in less than seven. [8]

A unique aspect of golf, as opposed to say tennis, is that each tour provides a set of golfers a *tour card* which grants membership status for a full year. A tour card provides a golfer the

[7] This may result from tournament exemptions provided to past champions which enable them to keep playing on the PGA TOUR despite a drop off in performance.

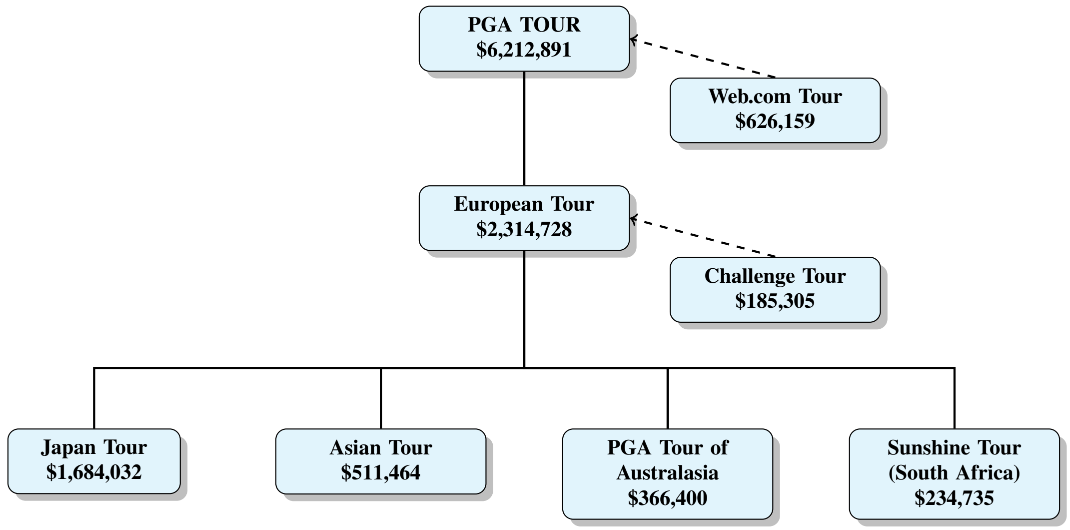
[8] These statistics are computed conditional on earning at least some money on the PGA TOUR in 2012.

opportunity to participate in most tournaments. Among those that hold a tour card, golfers are placed into categories designating the order in which they are eligible for tournaments. Some tournaments will have less open slots than there are players with tour cards. Therefore, of the players with tour cards that apply, only those with the best status are accepted. However, all players granted entry through either the Web.com Tour ML or Q School have the opportunity to play in many events.

In addition to the PGA TOUR, there are many golf tours throughout the world that offer substantial prize money. The vast majority of the golfers on these tours are professional and use their golf earnings as their primary or sole source of income. Figure 2.2 shows a hierarchy of golf tours in terms of the money awarded and quality of players. Golf tours are ordered from most prestigious (top) to least prestigious (bottom). The dollar figure represents the average purse in the 2012 season for each tour. The PGA TOUR offers the most lucrative prizes, awarding approximately three times the money as the European Tour. A dashed line denotes a relationship between a developmental tour and a major tour. For instance, the Web.com Tour is owned and operated by the PGA TOUR. The main purposes of the Web.com Tour are to prepare players to compete on the PGA TOUR and offer players that fail to attain membership a place to earn a living playing golf in hopes of making it back to the PGA TOUR later. Developmental tours pay substantially less than major tours (about one tenth of the prize money), but can pay enough to make a good living. For instance, in 2012, the Web.com money leader, Casey Wittenberg, earned $433,453. Out of 287 players who earned some money, 57 earned over $100,000 and 106 earned over $50,000.

Given the disparities in pay and prestige, it is generally safe to assume that all golfers hope to eventually play on either the PGA TOUR or the European Tour. As a result, success on a lower tour is generally accompanied by a transition to a higher tour. The converse is also true, poor performance on a major tour often results in a demotion to a lower tour. Prize money on second tier tours has been increasing, resulting in more US or European golfers choosing to play abroad.

Figure 2.2: Hierarchy of Professional Golf Tours



Note: Figure includes the average tournament purse in 2012 for each tour.

2.3.2 Earnings Data

Given that golf is an increasingly global game, it is important to collect worldwide earnings information. I compile earnings data from eight golf tours from around the world from 1985 to 2014.[9] To produce this dataset, I merge earnings of each player from every tour based on either an identification number or golfer name. The most important source of earnings data for this study comes from the PGA TOUR and Web.com Tour. Players on these tours can be linked through a unique identification number, providing a highly reliable merge. Earnings data from tours other than the PGA TOUR and the Web.com Tour are assembled through a merge on full name. Duplicate names are not common for a few reasons. First, the sample size of the earnings data is rather modest. From 1990-2012, there are 3,696 unique golfers with positive world earnings. Therefore, some corrections can be made through visual inspection. Second, golfers and golf organizations

[9] The Web.com Tour and the Challenge Tour started in 1990. Asian Tour and Sunshine Tour data being in 1995 and 1991, respectively. See Table 2.3 for the full details of earnings data availability by tour.

try to keep names unique. Since golf is an individual sport a player's name serves as something analogous to a brand or team name. If a player has the same or similar name another he will often add a middle name to differentiate himself. Therefore, failed merges will often be the result of spelling errors or structural data management differences between tours. For instance, some tours store the full middle name whereas some only include a middle initial. Spelling differences will reduce the success of the merge. However, as these errors are based on names, it is plausible that this measurement error lacks any systematic bias.

The key features of the data is that it is administrative and a long panel. Therefore, the quality is high and earings can be observed across a career on multiple tours. *World earnings* refers to aggregate earnings across all tours.

In order to make comparisons of earnings across time, two adjustments are required. First, a standard adjustment for inflation is appropriate. Second, and more unique to the golf setting, earnings should be adjusted for an increase in the demand for watching golf. As golf grew in popularity, prize money far outpaced the rate of inflation. For instance, prize money grew at a 12.5% annual rate in the 1990s and 7.8% in the 2000s.[10] Without an adjustment, the estimation of any statistics that pools data across time would implicitly weight more recent earnings observations more heavily. Therefore, I adjust for the rate of growth in prize money to account simultaneously for inflation and increased demand.

Specifically for each golfer i, I multiply total earnings, e, in year t by a normalization factor equal to the official PGA TOUR tournament average purse in 2012 divided by the same statistic in year t. Let m represent average purse, then the following formula shows the explicit normalization

[10] The rate of growth in prize money slowed to 2.3% from 2010 to 2014, likely related to the Great Recession and its aftermath.

so that all earnings are in terms of 2012 dollars

$$\tilde{e}_{i,t} = e_{i,t}\frac{m_{2012}}{m_t} \qquad \forall i, t. \tag{2.1}$$

Figure 2.14 shows the evolution of average PGA TOUR tournament purse over time.

Annual golf earnings is an aggregation of individual tournament prize money over the course of a season. The PGA TOUR awards prize money in a manner which is skewed towards the top performers. They use a payout system that is very consistent across tournaments and across time where the winner earns 18% of the total purse, second place 10.8%, third 6.8%, fourth 4.8%, and fifth 4%. Usually around 70 golfers will make the cut and thus earn some amount of money. Payoffs taper out such that 70$^{\text{th}}$ place earns just 0.2% of the purse. The payoff system is heavily skewed with 44.4% of the purse going to the top 5 finishers, 60.05% to the top 10, and 83.35% to the top 25.

Given the strong skewness in the prize money distribution, it is not surprising that annual earnings distributions are also highly skewed. Figure 2.3 compares the earnings distribution relative to a normal distribution for the Q School experiment. The blue dots represent earnings observations, the red line represents a normal distribution with the same mean and variance, and the gridlines represent the 5$^{\text{th}}$, 10$^{\text{th}}$, 25$^{\text{th}}$, 50$^{\text{th}}$, 75$^{\text{th}}$, 90$^{\text{th}}$, and 95$^{\text{th}}$ percentiles of each distribution. If earnings are normally distributed, the the blue dots will line up on top of the red line. Figure 2.3a plots the future earnings distribution and Figure 2.3b the future log earnings distribution. Figure 2.3a shows that the earnings distribution is a poor approximation for a normal distribution. There is a large clumping of observations near zero and many extreme positive outliers. Figure 2.3b, however, shows that the log earnings distribution is much closer to normally distributed with some slight divergence in the right tail.[11] Given the modest sample sizes of the experiments, it is impor-

[11] Figure 2.17 shows very similar results for the Web.com Tour ML experiment

tant to use log earnings, instead of earnings, since it closely approximates a normal distribution. However, using log earnings precludes a computation of the total effect on earnings–where total effect denotes the employment effect multiplied by the conditional earnings effect. In some cases, therefore, I refer the noisier level earnings estimates when addressing the total effect.

Earnings data are constructed for golfers between 17 and 55 years of age. I remove golfers that hold exemptions which potentially weaken the treatment effect. Exemptions include conditional status on the PGA TOUR, Web.com ML exemptions (for the Q School sample), medical exemptions, and career money list exemptions.[12] I discuss these exemptions in further detail in Section 2.4.2.

Earnings from tournament prize money is not the only source of income for professional golfers. The best players get sponsorship deals from firms which can often be substantial. However, lucrative contracts of this type are largely reserved for only a handful of the most recognizable players and sponsorships are likely to be highly positively correlated with performance. I have no data on sponsorship earnings.

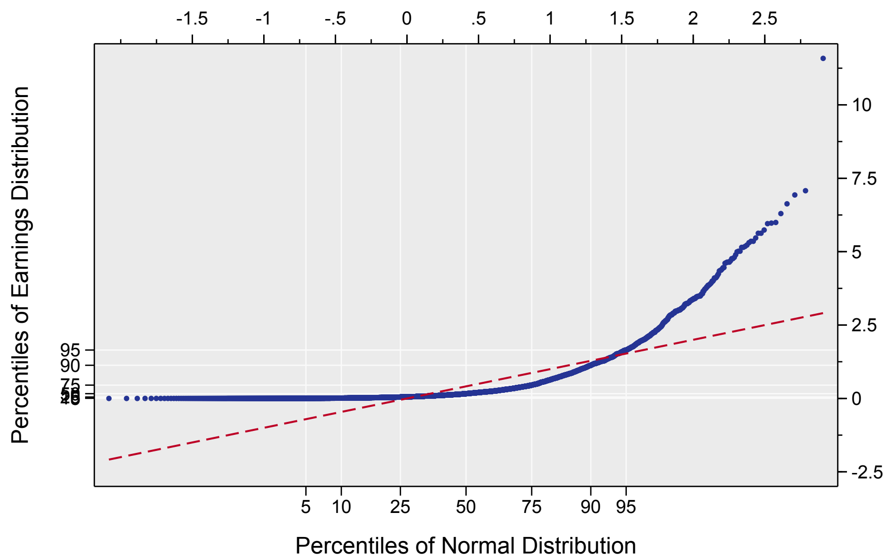
2.3.3 Official World Golf Ranking Data

I also compile data on Official World Golf Rankings (OWGR) which are produced by a collaborative organization of professional golf tours and rule making bodies. OWGR's are computed based on a system which award points to a golfer based on performance in the previous two years, adjusting for the prestige of golf tournaments played. I use the OWGR's as a proxy for ability as it is the best measure available. The main value of this measure is that it attempts to rank golfers who play on different tours throughout the world. Direct measures of performance are only reliable provided that golfers play in the same events. It is much more difficult, however, to rank golfers

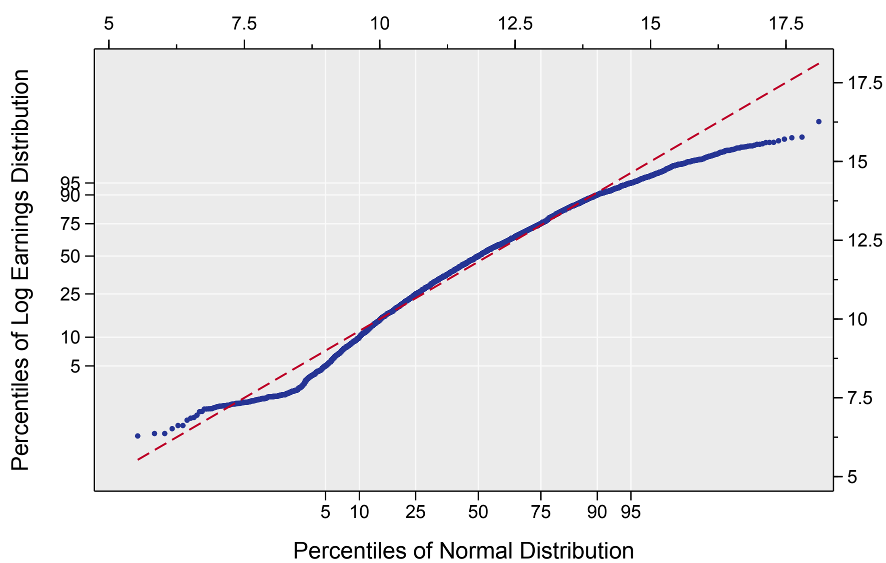
[12] Also, other potential exemptions are given for winning tournaments in recent years but no golfers hold these exemptions in the experiment samples.

Figure 2.3: Earnings Distributions Relative to a Normal Distribution for Q School Experiment

(a) Future Earnings Distribution



(b) Future Log Earnings Distribution



Notes: *Earnings from 1, 2, 3, 4, 5, 10, 15, and 20 years after treatment are aggregated and plotted. The left and bottom axes show the location of the notes percentiles of the distribution. The top and right axes show the scales. Earnings are in millions of dollars. Appendix Figure 2.17 shows very similar results for the Web.com Tour ML Experiment.*

when they play in very different events. Since OWGR's are produced through a collaboration of global tours and rule making bodies, it is reasonable to expect this measure to accurately judge the relative prestige of the tournaments on each tour. However, like all measures of ability based on past performance, OWGR is not a perfect proxy. For instance, young golfers who have played few events may have very high ability but will have few points and, hence, a poor ranking. Conversely, older or injured golfers may have low ability, but a high ranking from past performance.

I link OWGR data to earnings data through a merge on full name. For those observations that fail to merge I impute the highest possible OWGR for that month. In many cases this is the correct ranking as the player has in fact accumulated zero points. In other cases, there can be mistakes in the spelling of names which inhibit an accurate merge. In this case the imputation will produce measurement error. However, it is reasonable to expect this measurement error to not be systematically biased since it based on the random characteristics of a name.

2.3.4 Scoring Data

Finally, I have been granted access to scoring data for the PGA TOUR and the Web.com Tour from 1990-2014. The score in any given tournament is a direct measure of productivity. In fact earnings are awarded only on the basis of final score. Therefore, conditional on the quality of events entered, a golfer's annual scoring average is an excellent measure of annual performance. However, scores are only comparable at the same golf course at the same time. Golf courses vary in their difficulty and the difficulty of a particular course varies over time depending on weather conditions and course setup.

To resolve these issues I perform two adjustments. First, I subtract the field average score from each golfer's score in each round. This step removes a golf course/day fixed effect. However, this measure introduces a new dependency–the quality of the field.

Therefore, in a second step, I control for field quality in the following manner. Let i denote golfer, t year, and r round.[13] Also, let s denote score and x OWGR. Relative score, $(s_{itf} - \overline{s}_{tf})$, is a function of both individual ability and group ability. I use the OWGR of each golfer i in round r and year t as a proxy for individual ability and field average OWGR in round r and year t as a proxy for group ability. With a third order polynomial in both x_{it} and $\overline{x}_{rt}$, I nonparametrically approximate the relative score function. The following equation represents a score determination equation where the β parameters represent factors of individual ability and the γ parameters represent group ability:

$$(s_{irt} - \overline{s}_{rt}) = \alpha_t + \beta_{1t} x_{irt} + \beta_{2t} x_{irt}^2 + \beta_{3t} x_{irt}^3 + \gamma_{1t} \overline{x}_{rt} + \gamma_{2t} \overline{x}_{rt}^2 + \gamma_{3t} \overline{x}_{rt}^3 + \epsilon_{irt}. \tag{2.2}$$

I estimate equation (2.2) for each year separately. To adjust for field quality, I then insert the average OWGR for all rounds in year t, $\overline{\overline{x}}_t$, for $\overline{x}_{rt}$ as shown in the following formula:[14]

$$\widetilde{(s_{irt} - \overline{s}_{rt})} = \hat{\alpha}_t + \hat{\beta}_{1t} x_{irt} + \hat{\beta}_{2t} x_{irt}^2 + \hat{\beta}_{3t} x_{irt}^3 + \hat{\gamma}_{1t} \overline{\overline{x}}_t + \hat{\gamma}_{2t} \overline{\overline{x}}_t^2 + \hat{\gamma}_{3t} \overline{\overline{x}}_t^3 + e_{irt}. \tag{2.3}$$

The key feature of this adjustment is that it plausibly allows a comparison of scores in the years after treatment. Since the treated players mostly play in PGA TOUR events while control players mostly play in Web.com Tour events, the average strength of tournament field is systematically higher for those in treatment. Adjusting scores for the strength of field alleviates this problem.

[13]The round encompasses the golf course, the tournament, and the day in which the golf was played.

[14]It is important not to interact an individual golfer's OWGR with field average OWGR since an individual golfer's OWGR will change with treatment. This could create spurious treatment effects in the future adjusted scoring average.

2.4 Details of the Natural Experiments

2.4.1 PGA TOUR Qualifying Tournament

From 1962 to 2012, the PGA TOUR held an annual tournament called the PGA TOUR Qualifying Tournament or Q School.[15] The top finishers in this tournament were awarded a PGA TOUR card. Q School was historically the main method by which golfers earned a PGA TOUR card.

Q School is a four stage tournament with final stage, second stage, first stage, and pre-qualifying stage. The final stage consists of six rounds of golf. Stages two, one, and pre-qualifying have four, four, and three rounds, respectively. Depending on a player's status he may be exempt from any or all of the stages. Most regular tournaments only have four rounds. Given the number of stages and length of each stage, Q School is a grueling event. For example in 2005, 1205 players competed, but only 32 were awarded a tour card.[16]

In 2012 the final stage awarded the top 25 finishers plus ties a PGA TOUR card. The next 50 players plus ties earned a Web.com tour card. I refer to those that receive a PGA TOUR card as the *treated* group and those that receive a Web.com Tour card as the *control* or *untreated* group. A player who narrowly qualifies can expect to play in about 23 PGA TOUR events the next year, whereas a player who barely misses qualification can only expect to play in 2 events. The RD design is built on the assumption that ability is continuous across the treatment threshold. Continuous ability along with discontinuous benefits at the threshold provide an exogenous source of variation in employment states. One challenge intrinsic to this design is that many players tie around the threshold. The highly discrete nature of the running variable suggests using different

[15] Q School was discontinued in 2013. Tour cards are now given only to golfers on the top 50 of the season end Web.com Tour ML. The PGA TOUR website states: "The change was made due to a variety of factors, but the overwhelming success of Web.com Tour graduates over 20+ years on the PGA TOUR was the primary motivation for the change." <http://www.pgatour.com/company/pga-tour-faqs.html>

[16] The source of these numbers is Feinstein (2007) which is national bestselling book documenting the challenges of the Q School process.

methods from the standard continuity based local linear regressions. Section 2.5 discusses the details of the identification and estimation of the RD design.

Although the structure of the stages remained fairly constant since 1990, the number of tour cards granted at final stage decreased over time.[17] Figure 2.16 shows a decrease from 46 tours cards in 1993 to 25 in 2012. At the same time tour cards awarded from the Web.com Tour ML increased. The aggregate effect is that about the same number of golfers earned a tour card through both channels over time. The PGA TOUR increasingly emphasized access through the Web.com Tour ML as they perceived the quality of golfers to be of higher average ability with less variance (Feinstein, 2007). This feature may arise naturally since the money list is a longer-term measure of performance. If treatment has heterogeneous effects based on ability and the changing thresholds effect the average ability of players subject to treatment, then the measured effects of treatment will change over time. Appendix 2.11.2 reports the results of a sensitivity analysis where I estimate treatment effects with two samples split by time period. The results are fairly stable over time, alleviating these concerns.

Due to data limitations, the Q School experiment is estimated in years 1993 and 1994-2012. There are records of the successful qualifiers, but none remain for the golfers who failed to qualify through Q School final stage in 1990-1992 and 1994.

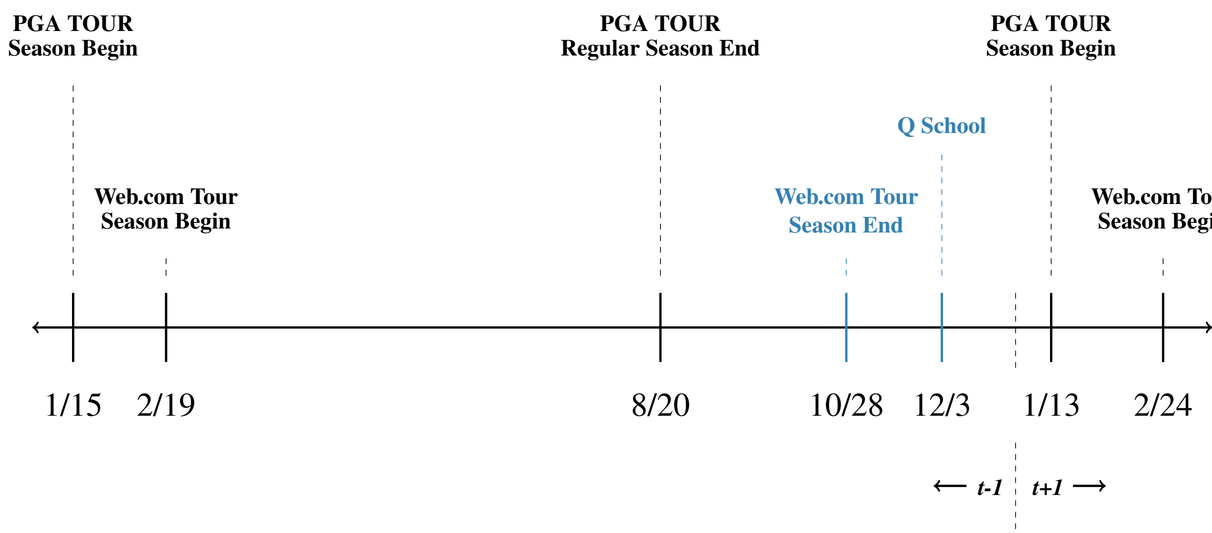
2.4.2 Web.com Tour Money List

In 1990, the PGA TOUR created a developmental tour called the Web.com Tour.[18] The goals of the Web.com Tour are to develop a competitive stock of golfers to continually enter the PGA TOUR

[17] One structural change occurred in 2006 when the pre-qualifying stage was introduced.

[18] At the time the tour was named the Ben Hogan Tour and at different times has been called the Nike Tour, the Buy.com Tour, and the Nationwide Tour, depending on sponsorship.

Figure 2.4: Time Line of Events (Example: 2012-2013)



and to provide some stability for those who fail to return to the PGA TOUR. The prize money awarded on the Web.com Tour is about one tenth of that awarded on the PGA TOUR.[19]

The Web.com Tour hosts about 29 tournaments a year. In each tournament golfers are awarded prize money based on their final position. The money list refers to cumulative year-end prize money. In 2012, golfers ranked in the top 25 on the money list earned a PGA TOUR card while those finishing between 26 and 50 earned a Web.com Tour card. Therefore, the treatment for this natural experiment is very to similar to that of the Q School experiment, where the treated receive a PGA TOUR card and the control, a Web.com Tour card. The last golfer in can expect to play in 25 PGA TOUR events the following year, while the last golfer out can expect to play in 8 events.

Once again the discontinuity in benefits across the treatment threshold and an assumption of continuous ability creates the conditions for identification in a RD design. The treatment is very similar to the Q School treatment. Each experiment should serve as a replication of the other. However, there are subtle differences in the treatments with respect to the running variables. First,

[19]Figure 2.15 shows the fluctions in Web.com Tour prize money to PGA TOUR prize money. Although fairly stable, the ratio has fluctation between 0.15 and 0.09.

Q School finish position is a shorter term measure of ability than Web.com Tour ML season rank. Therefore, we should expect it to be less predictive of future success. Second, in a given year golfers can tie at the same Q School finish position, whereas there are no ties in the Web.com Tour ML rank. Third, the Web.com Tour threshold affects golfers of slightly higher ability. I present this evidence in more detail in Section 2.4.3. Finally, the Web.com Tour golfers can attempt to subsequently earn a tour card through Q School, for which they are exempt into the final stage. Figure 2.4 presents an example time line of events for clarification. The opportunity to subsequently gain a PGA TOUR card through Q School slightly weakens the Web.com Tour ML treatment.

As previously noted I remove golfers from the sample with conditional status on the PGA TOUR, a Web.com Tour ML exemption (for the Q School sample), a medical exemption, or a career money list exemption. Figure 2.16 plots the relationship between qualifiers and exempt qualifiers for each experiment. Most exempt qualifiers are from the Q School experiment and have conditional status on the PGA TOUR as a result of finishing between 126 and 150 on the previous years PGA TOUR ML. Conditional status provides an exemption into the final round of Q School and PGA TOUR membership benefits at lower priority than Q School or Web.com Tour ML graduates. On average players with conditional status play in 20 PGA TOUR events in the next year, whereas Q School graduates play in 25 events. As a result many players with conditional status compete in Q School. Medical exemptions are given to players with a tour card in the previous year but missed many events due to a serious injury. Since the Web.com Tour season ends before Q School, many players choose to play in Q School after the Web.com Tour season. In fact, in 2012 those finishing between 26 and 40 on the Web.com Tour ML were exempt into final stage of Q School. On rare occasions, even players that earn a tour card by finishing in the top 25 on the Web.com Tour ML play in Q School. Within each status group players are placed in the following order: 1st place Q School, 2nd place Web.com Tour, 2nd place Q School, 3rd place Web.com Tour, ... , last place to qualify Q School, 25th place Web.com Tour. In practice this ordering makes little difference in the number of tournaments a player will have access to in the next year. However, this creates a possibility for players who earned a tour card through the Web.com Tour money list to improve their order in the queue through Q School qualification. Across the nineteen sample years, 27 golfers enter Q School after qualifying through the Web.com Tour ML.

A tour card provides a golfer with exemptions into many of next year's PGA TOUR events. However, there are alternative ways to qualify for PGA TOUR events. For instance, players can qualify for single events each week through Monday qualifiers–one round tournaments that accept only a few golfers. Also players can be offered special sponsor's exemptions to play in events. These aren't necessarily awarded based on merit, but only a handful are given. If a player finishes in the top 25 of a tournament, after say successfully Monday qualifying, he is exempt to play in next week's PGA TOUR event. Furthermore, if a player earns an amount of prize money greater than or equal to last year's 100th place finisher on the official PGA TOUR money list, then he is granted a tour card for the remainder of the season. Therefore, there are multiple paths to playing on the PGA TOUR, but without a tour card it is very difficult to sustain a presence.

2.4.3 Descriptive Statistics of Pre-Treatment Characteristics

Table 2.1 presents descriptive statistics of relevant pre-treatment characteristics from the Web.com ML and Q School samples. Summary statistics are computed from the 10 closest players to each side of the treatment threshold for both experiments separately for qualifiers and non-qualifiers. I break ties in the running variables by sorting golfers with the same score alphabetically by first name. These summary statistics do not incorporate a slope term for the running variable. Therefore, differences between qualifiers and non-qualifiers should not be interpreted as representing underlying pre-treatment differences across the treatment thresholds. Instead Table 2.1 is meant to present the differences across experiments in summary statistics of pre-treatment age, ability, and experience of golfers close to the treatment thresholds.

First, note that mean and median age of golfers in both experiments is very similar. Second, note that Web.com Tour ML golfers appear to be of higher ability. For instance, the mean and median OWGR are lower in the Web.com ML population than in the Q School sample. Also, average earnings in the past five years are higher in the Web.com ML population. Third, neither population is obviously more experienced than the other. The *past events* variables represent the total past events entered in all previous years on either the PGA TOUR or the Web.com Tour, respectively. The Q School sample has more experience on the PGA TOUR, but the Web.com Tour ML sample has more experience on the Web.com Tour. This is not surprising since in order to be in the Web.com Tour ML sample a golfer must have played at least one full season on the Web.com Tour.

Another clear result of Table 2.1 is that the Q School sample has a larger standard deviation for all pre-treatment characteristics except past Web.com Tour experience. This is consistent with expectations since the Q School experiment selects qualifiers based on a more short-term measure of performance than the Web.com ML experiment.

Table 2.1 provides important information for interpreting the results of Section 2.6 with the main takeaways being that the Web.com ML sample is of slightly higher ability and represents a more homogeneous population. If there are heterogeneous treatment effects with respect to ability, then the variance of the Q School sample could lead to higher variance in the estimation results.

Table 2.1: Age, Ability, and Experience Descriptive Stats

		Web.com ML		*Q School*	
		Qualifiers	Non-Qualifiers	Qualifiers	Non-Qualifiers
	Mean	31.5	31.9	31.3	32.3
Age	Median	31.0	31.0	31.0	31.0
	SD	5.4	5.7	6.2	6.4
	Mean	485.0	558.9	730.0	774.0
OWGR	Median	440.0	517.0	740.0	755.0
	SD	200.5	194.8	353.2	350.1
	Mean	12.05	11.88	11.57	11.43
Past Log World Earnings	Median	12.05	12.02	11.89	11.72
	SD	0.82	0.80	1.59	1.52
	Mean	51	52	69	72
Past PGA TOUR Events	Median	28	10	14	13
	SD	76	86	107	117
	Mean	80	84	47	46
Past Web.com Tour Events	Median	66	71	27	29
	SD	56	58	55	52

Notes: *Sample comprised of nearest 10 golfers on both sides of the treatment threshold. Sample size is 190 for all statistics as only years 1993, 1995-2012 are used for both experiments. "Past Earnings" denotes average earnings in the previous 5 years. "Past Events" refers to total past events.*

2.5 Identification and Estimation

The RD designs are based on two general assumptions. First, treatment is given to all golfers below a cutoff and withheld from all above the cutoff. Second, golfers in the control and treatment groups near the cutoff are valid counterfactuals for each other, which is to say that they cannot precisely manipulate their score value and hence their treatment status.

The first assumption applies to the case of a *sharp* RD design where compliance is either perfect or attention is focused on the intention-to-treat parameter. Employing a sharp RD design, I implicitly define the treatment to be either qualification through the Web.com Tour ML or Q School, rather than, for instance, a year's membership on the PGA TOUR. An issue with the latter definition of treatment is that it requires an arbitrary and binary definition of PGA TOUR membership. For example, suppose I define all golfers who play in at least 22 PGA TOUR events to be PGA TOUR members. With this definition I could employ a *fuzzy* RD design which essentially involves scaling the estimated coefficients by the differential probability of treatment at the cutoff. This design speaks more directly to the question of how a year of PGA TOUR membership affects career outcomes. However, the weakness in this strategy lies in the arbitrary and binary definition of PGA TOUR membership. Given that the results would be sensitive to this definition, I prefer, for reasons of transparency, to show the sharp, intention-to-treat results. The reader can then consider the effects on future earnings and employment jointly, keeping in mind the relative strengths of the two experiments.

The second assumption is formalized in Sections 2.5.1 and 2.5.2 and tested in Section 2.5.4. This is the key identifying assumption of the local average treatment effect for the RD design. For what follows define an outcome of interest as Y_i, a running variable (or score) as R_i, and treatment status as D_i. The experiments have different running variables. In the Web.com Tour ML experiment, the running variable is a more continuous measure, the final position on the Web.com Tour ML. In the Q School experiment, the running variable is more discrete, the final score in the final stage of Q School. By discrete, I am specifically referring to the likelihood of observing multiple observations at the same value of the running variable. These differences in the nature of the running variables warrant the use of different identification assumptions and estimation methods.

In Q School experiment, in each year there are many golfers that share a final score, and this is especially true around the treatment cutoff. The distribution becomes even more discrete when experiments are stacked to create one large experiment. Figure 2.7a shows 109 treated and 121 control golfers within one stroke of the cutoff.

In contrast, there are no ties on the Web.com Tour year-end money list. As a result within each individual year, the running variable is continuous. However, as experiments are stacked to create one large experiment, the running variable becomes more discrete. Figure 2.7b shows that there are 23 golfers on either side of the Web.com Tour ML. Although discrete in the sense that multiple golfers hold the same value, this running variable is closer to continuous, and therefore, the Web.com Tour ML experiment can be approximated with continuous methods.

In the continuous case, I employ the standard RD assumption that the conditional expectations of potential outcomes given the score are continuous at the cutoff, ensuring that the characteristics of golfers do not abruptly change across the treatment threshold (Hahn et al., 2001). In the discrete case, I employ *local randomization* methods where the treatment is assumed to be randomly assigned in a small window around the cutoff (Cattaneo et al., 2015). I discuss these assumptions and their corresponding estimation methods below.

2.5.1 Continuous Running Variable

The Web.com Tour ML experiment has a running variable, final money list position, that is approximately continuous. As a result the primary analysis is performed with continuous running variable methods, while a secondary results are shown with discrete methods in the appendix for robustness.

Adopting the potential outcomes framework with a sharp RD design, each individual with observed running variable R_i above the known threshold $\overline{r}$ is assigned to the control group ($D_i = 0$) and each individual with R_i below $\overline{r}$ is assigned to the treatment group ($D_i = 1$).[20] Thus, $D_i = \mathbb{1}(R_i \leq \overline{r})$ for each individual i in the sample. Each individual has two potential outcomes, $Y_i\,(1)$ and $Y_i\,(0)$, corresponding to treatment status. Thus, the observed outcome is:

$$Y_i = Y_i\,(0) \cdot (1 - D_i) + Y_i\,(1) \cdot D_i. \tag{2.4}$$

In this application, $Y_i\,(0)$ represents future earnings, events, employment, or scoring average in the absence of qualification for a PGA TOUR card through Q School or Web.com Tour ML, but instead, receiving a Web.com Tour card. $Y_i\,(1)$ captures future outcomes if the golfer receives a PGA TOUR card through Q School or the Web.com Tour ML. The main parameter of interest is the population average response to treatment at the cutoff $\overline{r}$:

$$\tau_c = \mathbb{E}[Y_i(1) - Y_i(0)|R_i = \overline{r}]. \tag{2.5}$$

Identification of τ_c, where c denotes a continuous running variable, requires the following assumption.

Assumption 1 (Nonparametric Continuous Identification)**.** $\mathbb{E}[Y_i(0)|R_i = \overline{r}]$ *and* $\mathbb{E}[Y_i(1)|R_i = \overline{r}]$ *are three times continuously differentiable at the RD cutoff* $r = \overline{r}$.

Utilizing this relatively weak assumption, I estimate the limits of the conditional expectations function on both sides of the cutoff. The bandwidths around these limits are key parameters for es-

[20] The cutoff, $\overline{r}$ is defined as the midpoint between the goflers than barely made and missed treatment. For instance, if the 25[th] placed golfer recieve treatment and the 26[th] placed gofler does not, then $\overline{r} = 25.5$.

timating the limits nonparametrically. I follow Cattaneo et al. (2016c) in implementing the following steps.[21] First, I choose a polynomial of degree one and a triangular kernel weighting function to compute the coverage error optimal RD bandwidths separately on each side of the cutoff, h_l and h_r. Second, I drop all observations outside the neighborhood $W = [\overline{r} - h_l, \overline{r} + h_r]$. Third, using weighted least-squares regression, I estimate a local linear regression allowing for different slopes on each side of the cutoff:

$$Y_i \quad = \quad \alpha + \tau D_i + \beta_1 \overline{R}_i + \lambda_1 \overline{R}_i D_i + \epsilon_i \tag{2.6}$$

with weights $\kappa\left(\overline{R}_i/h\right)$, and where $\overline{R}_i = R_i - \overline{r}$ is the re-centered running variable. The weighted least-squares estimates of τ in the above regression model, denoted by $\hat{\tau}$, estimates τ_c.

Calonico et al. (2014) note that methods for computing optimal bandwidths include a first order bias term resulting from bandwidths that are too "large". My main specification incorporates a bias correction to treatment effects and a standard error correction to compute robust confidence intervals. Robustness analysis reveals that these corrections do not change the qualitative conclusions. Additionally, Calonico et al. (2016) develop formal local polynomial methods allowing for pre-intervention covariate adjustments which can reduce standard errors. I adjust for age, age squared, and OWGR. I do not incorporate all pre-treatment characteristics in order to maximize sample size. When applying continuous running variables methods to discrete cases, Lee and Card (2008) recommend using standard errors clustered by the score to adjust for specification error. Since the stacked Web.com Tour ML position is not fully continuous, I cluster the standard errors by the running variable.

[21] I use the Stata software program *rdrobust* described in Calonico et al. (2017).

2.5.2 Discrete Running Variable

Given the highly discrete nature of finish position in the Q School experiment, methods designed for continuous running variables may not closely approximate the true average treatment effects (ATE's). Therefore, I employ the *local randomization* method proposed by Cattaneo et al. (2015) and Cattaneo et al. (2016c). In its most basic form, the idea behind local randomization is very simple. This approach assumes that, in a small neighborhood or window around the RD cutoff, the assignment of units to treatment or control status is random, as it would be in an experiment. Define the average treatment effect in the discrete setting as:

$$\tau_d = \mathbb{E}[Y_i(1) - Y_i(0) | i \in W_0]. \tag{2.7}$$

Identification of τ_d, requires the following assumption.

Assumption 2 (Local Randomization Identification)**.** *There exists window* $W_0 = [\bar{r} - w, \bar{r} + w]$*,* $w > 0$*, such that the score is independent of potential outcomes within* W_0*.*

A simple difference-in-means statistic across treatment status can be used to obtain an unbiased point estimator of ATE. Under the null hypothesis that the ATE is zero, an exact p-value can be computed for the difference in means statistic.[22] This allows the use of finite-sample exact randomization inference methods, where the null distribution of the test statistic of interest is derived directly from the randomization distribution of the treatment assignment inside the window, leading to inferences that are exact in finite samples.

The local randomization approach can work well in setting with a highly discrete running

[22]There are a few caveats to this identification arugment that must be addressed. First, the treatment assignment mechanism must be known within the window. Second, the treatment effect must be non-negative for all individuals, meaning $y_i(1) \geq y_i(0)$ for all observations in the window, to identify the ATE instead of the *sharp null hypothesis of no effect*. The stable unit treatment value assumption (SUTVA) is also required. See Cattaneo et al. (2015) and Cattaneo et al. (2016c) for the details of the technical issues invovled.

variable. The downside of this approach is that the randomization window must be known or estimated. However, in the Q School experiment there are enough observations within one stroke of the cutoff to power the estimation. Therefore, I select a window of within one stroke of the cutoff. This amounts to assuming that conditional on finishing within one stroke of the cutoff, the side a player ends up on is random. I then simply compare the difference in mean outcomes on each side of the cutoff. I assume a fixed-margin randomization scheme. Despite using such as small window, the inference methods formulated in Cattaneo et al. (2015) are able to deliver precise estimates.[23] For robustness, I employ the continuous running variable method to the Q School experiment. The results, shown in the appendix, are qualitatively similar, but have larger variance.

2.5.3 Plots of the RD Designs

Prior to estimating the RD effects, it is informative to plot the data around the cutoffs. I follow Calonico et al. (2015) in the implementation of my RD plots.[24] These plots have two main features. First, the plots include a global polynomial of an outcome variable with respect to the running variable in each experiment. The purpose of the polynomial is to represent the behavior of the underlying conditional expectations function in a smooth fashion and from a global perspective. Second, the plots include local sample means of the outcome variable constructed by partitioning the support of the running variable into disjoint bins separately for control and treatment golfers. The local means are meant to serve two purposes. First, they provide information regarding the validity of the identifying assumption of continuity of the conditional expectations function at the cutoff by allowing the reader to check for the presence of discontinuities away from the cutoff. A

[23] I implement the local randomization estimation using the Stata command *rdlocrand* which is described in Cattaneo et al. (2017).

[24] I use the Stata command *rdplots* described in Calonico et al. (2017).

second aim of the local sample means is to construct an under-smoothed estimate of the conditional expectations functions that displays the overall variability of the data.

Calonico et al. (2015) develop methods to compute sample means based on bin sizes that are optimal either for detecting discontinuities or representing the underlying variability of the data. I select bins such that the RD plots represent the underlying variability of the data. Modest sample sizes, particularly for treated golfers, represent a challenge to RD designs. Therefore, it is information to view a representation of that variation. The local sample means are constructed such that each bin has approximately the same number of observations on each respective side of the cutoff.

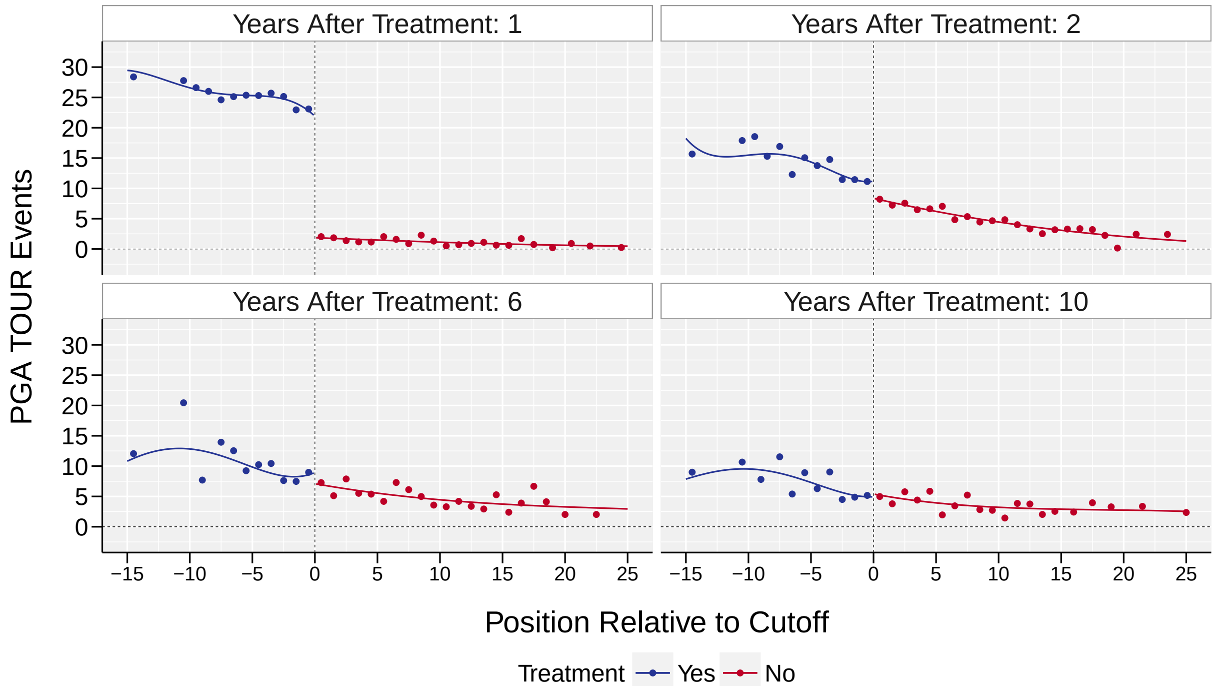
Figures 2.5 and 2.6 display some RD plots for both experiments, which also serve as a preview of the estimation results.[25] The first panel of figure 2.5 displays the average number of PGA TOUR events played in the first year after treatment for golfers on both sides of the cutoff. Clearly the treatments have a large effect in both experiments. Based on visual inspect, the qualifiers from Q School appear to play in 22.5 events in year one, whereas, the non-qualifiers only play in 2.5 events, for a difference of approximately 20 events. In the case of the Web.com Tour ML experiment, the qualifiers appear to play in about 25 events, whereas the non-qualifiers play in about 7.5 events, for difference of approximately 17.5 events. Both experiments appear to have a significant short-term impact with the Q School treatment being slightly stronger, as expected.

The remaining panels of Figure 2.5 show the number of PGA TOUR events golfers played in two, six, and ten years after treatment. Visually, the evidence appears weak for any persistent effects of treatment after the first year. In fact, already in year two the mean number of events played on each side of the cutoff appears to converge in both experiments. Note that particularly for the Web.com Tour ML experiment, at six and ten years after treatment the variance for treated

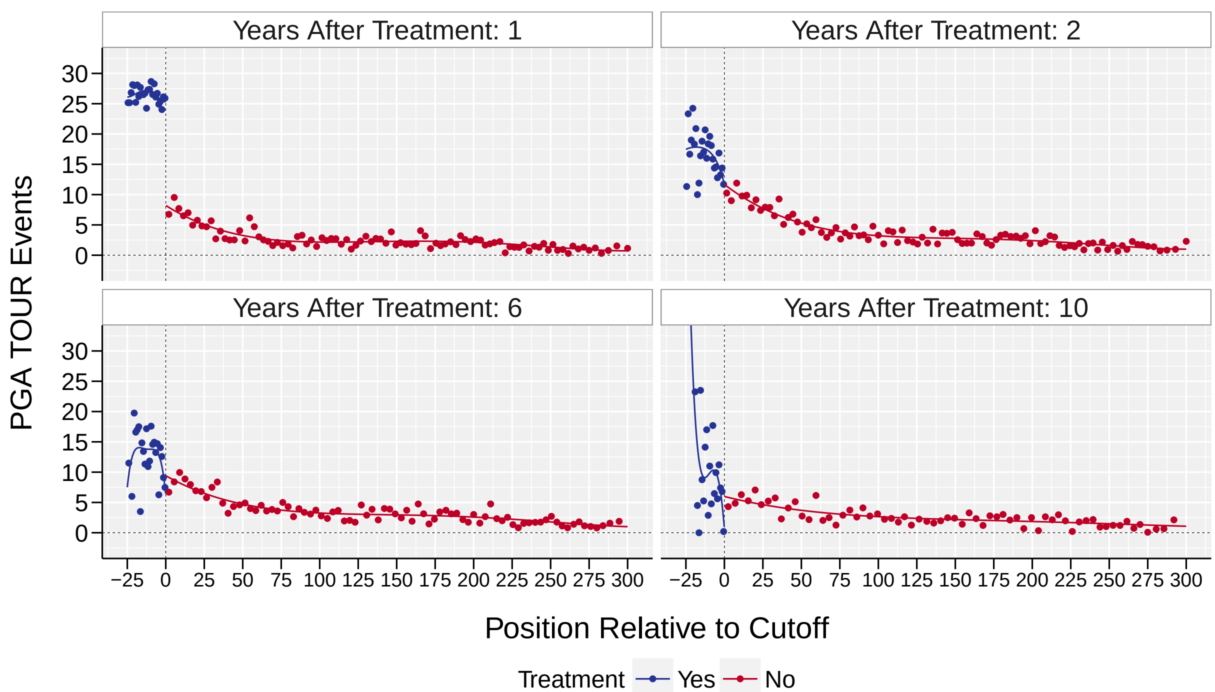
[25]However, one should bear in mind that the RD plots are meant to help the reader get an idea of the general features of the RD design, but the limits of the global polynomial at the treatment thresholds do not necessarily represent the limits estimated with the more formal estimation procedure.

Figure 2.5: RD Plots of Future PGA TOUR Events

(a) Q School



(b) Web.com Tour ML

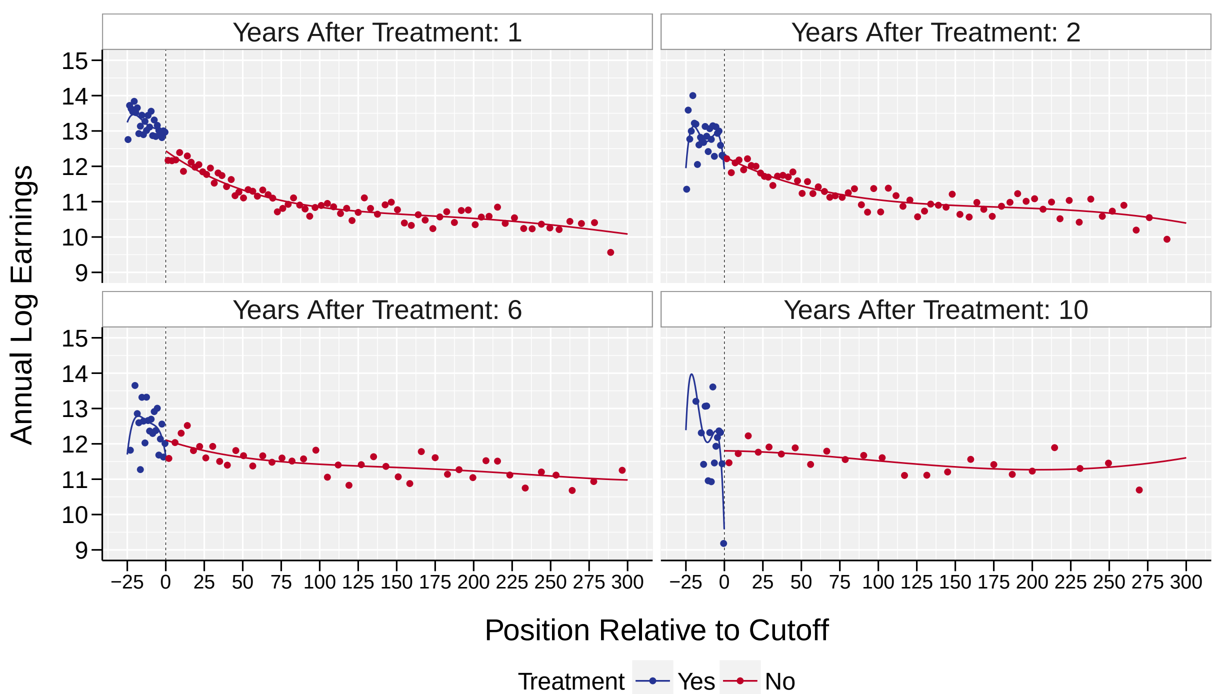


Notes: *Curves represent a global fourth degree polynomial fitted on each side of the cutoff. Dots represent averages within bins. Bins are selected optimally to represent the underlying variability of the data. Each bin is selected to have approximately the same number observations as other bins on the same side of the treatment threshold. The running variable for Q School is strokes from the cutoff, whereas the running variable for the Web.com Tour ML is money list positions from the cutoff.*

Figure 2.6: RD Plots of Future World Log Earnings

(a) Q School

(b) Web.com Tour ML



Notes: *See notes to Figure 2.5.*

players increases. Initially there were very few qualifiers from the Web.com Tour ML–only five in 1990 and 1991 (see Figure 2.16). This number has risen over time, but small samples in the early years reduce the number of observations with which to estimate long-term outcomes for the Web.com Tour ML experiment.

Figures 2.6 presents an RD plot of the main outcome of interest, future log earnings. In the first panel of both experiments, there appears to be a significant first year effect. However, the remaining panels do not appear to show any lingering effects. Also note how the sample variation grows as we look further into the future for both experiments. Furthermore, note that there does not appear to be any significant discontinuities away from the treatment thresholds for both outcomes in both experiments.

2.5.4 Validation of Regression Discontinuity Design

Assumptions 1 and 2 provide the conditions for identification of the ATE in each experiment. Assumption 1 applies to the Web.com Tour ML experiment and involves continuity in the conditional expectations function at the treatment threshold. Assumption 2 applies to the Q School experiment and is stronger, assuming randomization within some local window. Both of these conditions can be interpreted as statements about the ability of golfers to precisely manipulate their score and hence their treatment status. I argue that these RD designs do indeed satisfy the identification assumptions in two stages. First, I provide some details of the golf context to give the reader an idea of the feasibility of manipulation. Second, I supply the traditional data driven evidence of balance of observable pre-treatment characteristics and densities around the cutoffs. I conclude that these settings provide compelling conditions to employ RD designs and hence the results can be given a causal interpretation.

It is helpful to address the question of what manipulation would mean in this context. Two possible explanations come to mind. First, high ability players may exert low effort until the end of Q School or the Web.com Tour season, at which point they then may exert more effort and perform well enough to make the threshold. Under another scenario, suppose golf is played in two general environments: low and high pressure. Further suppose two golfers are of equal ability in low-pressure environments but one is better in high-pressure environments. Throughout the low-pressure, early stages of Q School or the Web.com Tour season, the golfers perform equally. Further suppose that a substantial proportion of play is under a low-pressure environment and, hence, their final standings are similar. However, at the end of Q School or the Web.com Tour season the good high-pressure golfer narrowly makes the cutoff and the low-pressure golfer misses. Suppose further that being good in high-pressure situations causes future PGA TOUR success– which is plausible given the skewed prize money allocation system.

The first scenario requires both foresight as to what the eventual threshold will be and enough control over outcomes to improve one's position at the right moment. The second scenario also requires good high-pressure players to have substantial control over their final score.

Whether or not golfers in Q School or the Web.com Tour are able predict the threshold with enough time to do something about it is debatable. Perhaps experienced players can make accurate guesses. However, given the inherent variability of golf, I do not expect players to be able to precisely manipulate their score on a week-by-week or day-by-day basis. Below I present some evidence on the inherent short-term volatility of golf performances.

On a year-to-year basis, past performance is a strong predictor of future performance. Indeed, the one-year correlation between year-end rank on the PGA TOUR from 2005-2014 is 0.73.[26] However, from tournament-to-tournament or round-to-round there is considerably more noise. Since golf is played outdoors, it is subject to the random effects of wind, weather, and the bounce of a ball on natural surfaces. These random elements can induce a large degree of short-term

[26]This is different from the 0.18 correlation number cited previoulsy as the sample consists of all PGA TOUR golfers rather than those ranked between 101 and 125 on the season ending money list. As the range of golfers is expanded, the correlation increases as there are greater differences in underlying ability.

fluctuations. In part in recognition of this fact, golf organizations structure tournaments to award champions based on a long-term average score such that each tournament is at least 72 holes. Indeed, I find that the correlation in golfers' finish positions between consecutive tournaments is only 0.01 on the PGA TOUR from 2005-2014. The correlation in scores from round-to-round within the same tournament is slightly higher at 0.13. These low correlations in performance at the tournament and round level provide quantitative evidence of the high short-term volatility inherent in golf.

Covariate Balance

In order to test the identification assumptions of continuity of the conditional expectation function at the cutoff (Assumption 1) for the Web.com Tour ML experiment and local randomization within a small window (Assumption 2) for the Q School experiment, I apply the appropriate RD estimation method to pre-treatment characteristics in each experiment. Table 2.2 displays the results of these regressions. The third column shows the estimates of τ_c and τ_d, without pre-estimation adjustments for other covariates. The fourth and fifth columns show the standard errors and p-values. The sixth and seventh columns show the number of observations and optimal bandwidth used to estimate the ATE on both sides of the cutoff. For all variables except age, I take an average from the previous five years to remove variance and increase precision.

For the both experiments, no pre-treatment characteristic shows a significant discontinuity at the treatment threshold at the 10% level. Overall there appear to be no significant differences in age, OWGR, experience, past earnings, or past scoring average in either the Q School or the Web.com ML experiments adding credibility to the identification assumptions.

OWGR for the Web.com Tour ML experiment is almost significant at the 10% level with a p-value of 0.103. However, given that I am testing eight covariates in two experiment for a total of sixteen tests, adjustments for multiple hypotheses are appropriate. If the pre-treatment characteristics are independent of each other, we would expect to see a significant ATE in one out of ten trials at the 10% level purely by chance. In this case with 16 separate outcomes, we expect to see 1.6 covariates with a significant ATE by chance. However, these characteristics are not independent from one another. Earning more money is correlated with playing in more events and having a lower scoring average. Positive correlation among covariates will reduce the expected number false positives, but should be approximately consistent with one covariate in sixteen showing mild significance.

Table 2.2: Covariate Balance Regressions

Experiment	Covariate	τ	se	p-val	$N_l \| N_r$	$h_l \| h_r$
Q School	Age	0.654	-	0.399	109 | 121	-0.5 | 0.5
	OWGR	4.138	-	0.931	109 | 121	-0.5 | 0.5
	World Earnings	-0.117	-	0.517	99 | 107	-0.5 | 0.5
	PGA TOUR Earnings	-0.074	-	0.713	99 | 107	-0.5 | 0.5
	Web.com Earnings	-0.096	-	0.642	99 | 107	-0.5 | 0.5
	PGA TOUR Events	-0.545	-	0.652	109 | 121	-0.5 | 0.5
	Web.com Events	0.424	-	0.580	109 | 121	-0.5 | 0.5
	PGA TOUR + Web.com SA	0.180	-	0.278	101 | 114	-0.5 | 0.5
Web.com ML	Age	0.626	1.325	0.637	299 | 974	-14.6 | 42.8
	OWGR	-62.758	38.450	0.103	319 | 1134	-16.7 | 50.2
	World Earnings	0.062	0.082	0.451	319 | 1046	-16.6 | 45.8
	PGA TOUR Earnings	0.067	0.092	0.466	319 | 1069	-17.1 | 47.0
	Web.com Earnings	0.087	0.107	0.415	319 | 1134	-17.0 | 50.4
	PGA TOUR Events	1.217	0.863	0.159	299 | 1112	-15.4 | 48.9
	Web.com Events	0.861	0.650	0.185	299 | 913	-15.4 | 39.8
	PGA TOUR + Web.com SA	-0.068	0.086	0.432	329 | 1112	-17.7 | 49.1

Notes: *All covariates except age are averages of the previous five years. Earnings are in log units. Q School effects are estimated with local randomization methods using the difference-in-means statistic. Web.com Tour ML are estimated with a local linear regression. See Section 2.5.1 for details. The local randomization software does not compute a standard error for the ATE.*

Densities at the Thresholds

Another common test of manipulation in a RD design is a test for a discontinuity in the density of the running variable across the treatment threshold. In both natural experiments, the number of qualifying golfers is determined prior to treatment. Therefore, I do not expect any discontinuities in the density of golfers around the thresholds, but still perform the tests. Also, plots of the densities help to understand the nature of the experiments. For each sample I remove "exempt" golfers from both treatment and control groups. Figure 2.7 shows histograms of each experiment by their respective running variables.

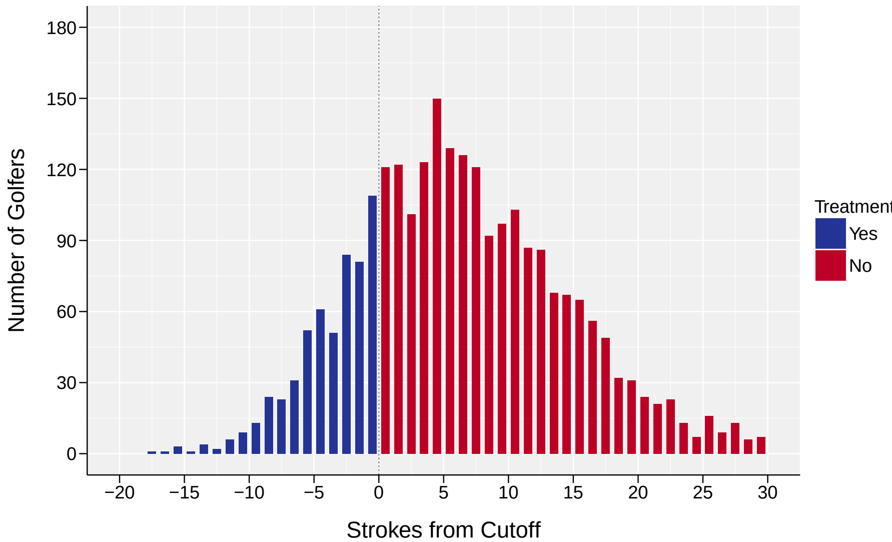
Figure 2.7a shows an approximate normal distribution for the Q School experiment with a slight positive skew. This is the aggregate distribution of Q School scores normalized by the threshold score in each year. There are many golfers within one shot of the threshold–109 treated and 121 control. Visual inspection shows no evidence of a discontinuity in the density at the cutoff. To formally test for a discontinuity in density at the cutoff I use methods introduced by Cattaneo et al. (2016b). I reject the null hypothesis of a discontinuity at the cutoff with a p-value of 0.34.[27]

Figure 2.7b displays the analogous histogram for the Web.com Tour ML experiment, but shows a very different distribution. Since the running variable is a year-long aggregate of prize money, there are no ties within a given year. I estimate the effects of the Web.com Tour ML threshold for 23 years (1990-2012). As a result for most scores there are 23 values. However, since I drop exempt players from my estimation sample, there need not be 23 by definition. Visual inspection shows no evidence of a discontinuity in the density at the cutoff. I reject the null hypothesis of a discontinuity at the cutoff with a p-value of 0.98.

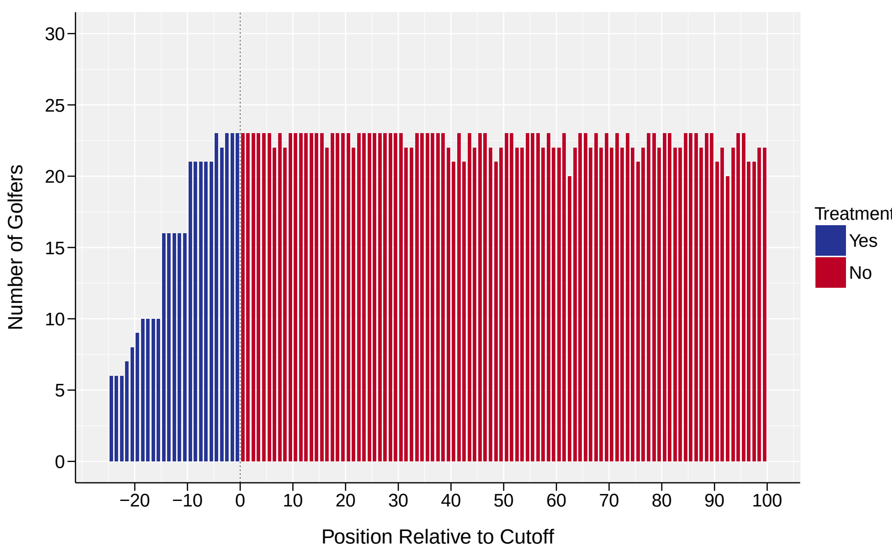
[27] I conduct this test using the software program *rddensity* described in Cattaneo et al. (2016a). I use default values in the unrestricted method which allows different MSE-optimal bandwidths on each side of the cutoff. This method is an improvement on the orginal density test suggested by McCrary (2008) as it does not require pre-binning of density values and it includes bias correction terms for optimal MSE bandwidths.

Figure 2.7: Histograms by Running Variable

(a) Q School



(b) Web.com Tour ML



Notes: *Histograms are computed for all "non-exempt" golfers in each respective sample. The Web.com Tour ML histogram trims observations for positions greater than 100 for a better view of the density around the treatment threshold.*

2.6 Results

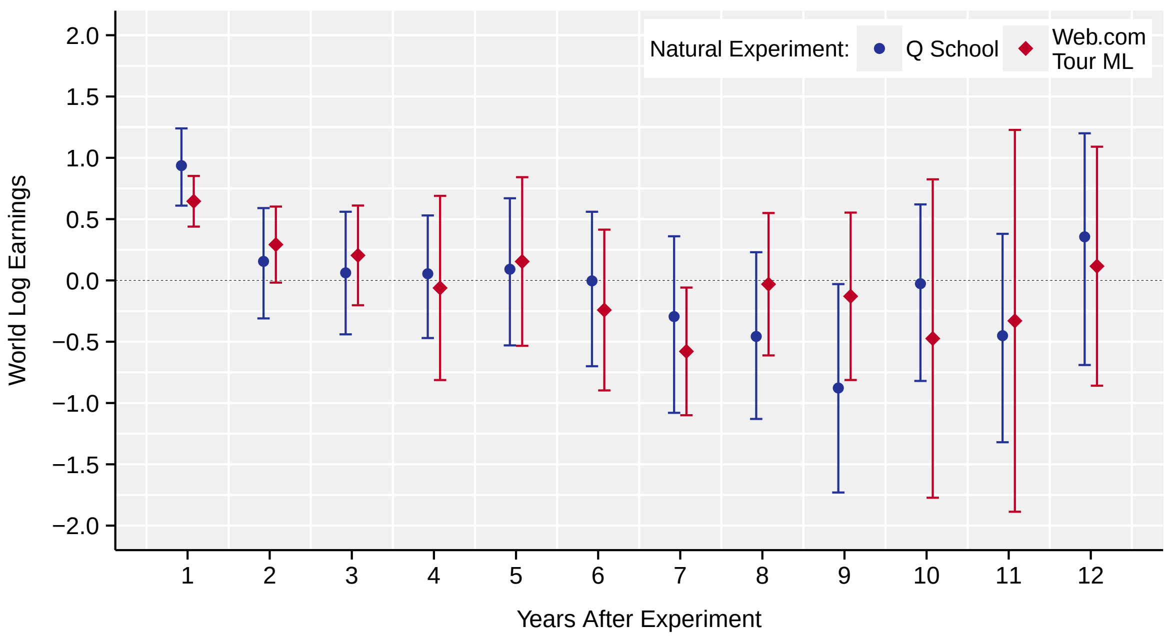
The main outcomes of interest are future earnings, productivity, and employment. In turn, I discuss earnings and productivity results and then employment results. The figures throughout Section 2.6 present the treatment effects in an event study format–showing effects one to twelve years after treatment. The circles and triangles represent the point estimates of the average treatment effect for each natural experiment. The bands represent 95% confidence intervals around these estimates. The appendix presents the exact values of the results in table format out to sixteen years after treatment.

2.6.1 Treatment Effects on Earnings and Productivity

Figure 2.8 presents average treatment effects for future log world earnings. Earnings effects are initially quite large. For the Q School experiment, I estimate an ATE on world earnings of 0.94 log points, or an approximate 94% earnings increase, in the first year. The analogous number for the Web.com treatment is 0.64 log points–a slightly weaker ATE but still a large effect. For perspective, job displacement studies using quarterly or annual administrative data in the US on average find an initial 30% loss in earnings (von Wachter, 2015). However, in contrast to job displacement studies where the average loss is 15% of earnings after 4 to 10 years, these effects do not persist. In both experiments, earnings differences quickly dissipate.

Despite the large initial effect of the Q School treatment, the earnings effect in year two is close to zero. The point estimate for the ATE is 0.16 log earnings points, but the p-value under the null hypothesis of a zero effect is 0.51. Therefore, due to a large variance, the earnings effect is statistically insignificant at any conventional confidence level for the second year after treatment. Effects in years three through six are consistently close to zero. In years seven through nine, there is weak evidence of negative treatment effects. However, there is substantial variance in the

Figure 2.8: Treatment Effects on Future World Earnings



Notes: *Circles and diamonds represent point estimates of the ATE's. Bands represent 95% confidence intervals. Q School is estimated with local randomization methods using the difference-in-means statistic with a window of one stroke around the treatment threshold. Web.com Tour ML is estimated with a bias corrected local linear regression. See Section 2.5.1 for estimation details.*

estimates in this range. Also, Section 2.6.2 documents positive employment effects in this range which largely offset the potential negative earnings effects. For the remainder of the observation period, the estimated effects hover around zero with large variance. This general pattern persists through years thirteen to sixteen (not pictured, see Table 2.10).

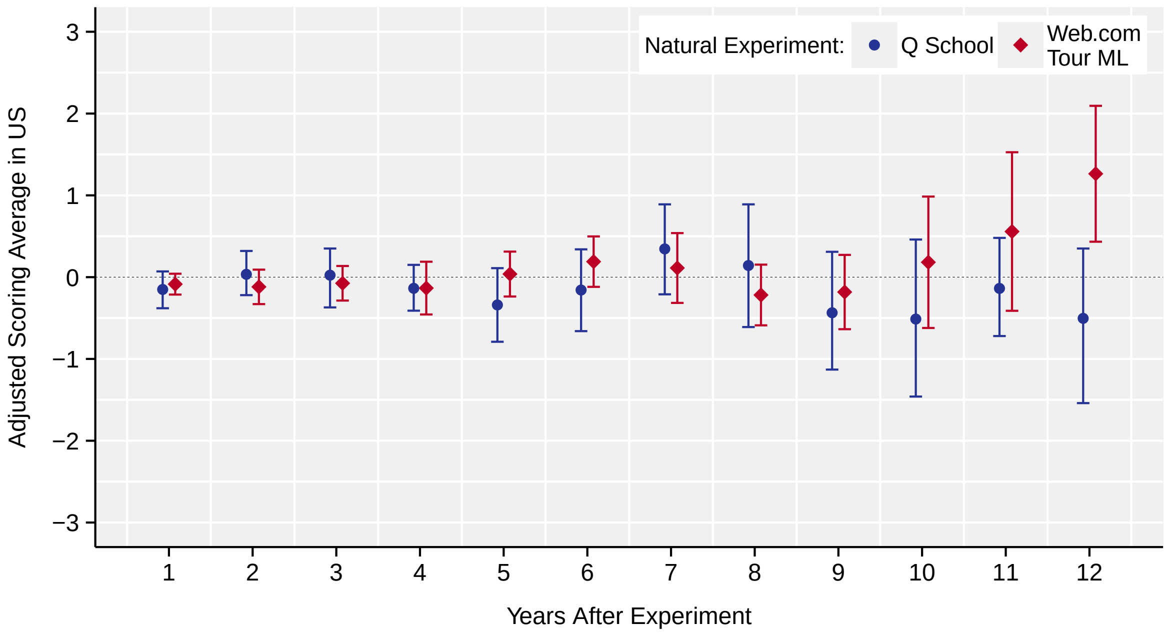
The Web.com Tour ML experiment shows broadly similar earnings results. However, despite an initially smaller treatment effect, there is more evidence of short-term earnings persistence. In years two and three, point estimates of the ATE's are 0.29 and 0.20 log earnings points with p-values of 0.065 and 0.325, respectively. Hence, there is some evidence that the treatment affects earnings at least in the second year. The sightly more persistent Web.com Tour ML effects could be the result of the higher average ability golfers in the sample. Perhaps these golfers are better positioned to take advantage of a lucky break.

Relative to the effects of employment shocks in the broader labor market, however, the consequences of the Web.com Tour ML experiment are more transitory. For the remainder of the observation period, no year has a statistically significant positive effect and the estimated effects generally hover around zero. There is a statistically significant, and quite large, negative effect in year seven. However, given that this effect is not consistent with the effects of neighboring years and the statistical significance of this effect is not robust across estimation methods, I conclude that this is more of a statistical artifact than a real effect.[28]

A feature of both experiments is the growing variance of the estimates as time passes. This is particularly true of the Web.com Tour ML experiment which is likely the result of the small number of treated golfers in the early years of the experiment. The Q School estimates have larger initial variance. This is likely due to either the discrete nature of the running variable or the greater variance of the sample in terms of ability. Although the estimated variances are large for both

[28]The seventh year effect is not robust to a local second degree polynomial regression specification (Figure 2.32) or the local randomization specification (Figure 2.34).

Figure 2.9: Treatment Effects on Future Scoring Average



Notes: *See notes to Figure 2.8.*

experiments, the similarly of the estimated effects across experiments provides some robustnesses to the results.

Figure 2.9 shows the estimated treatment effects on future adjusted scoring average–a measure of performance or productivity. Keep in mind that a low score is good. Therefore, an improvement in performance is represented by a negative treatment effect. In contrast to the earnings results, there are no initial treatment effects in either experiment. Within the sample of PGA TOUR and Web.com Tour golfers from 1995 to 2012, the standard deviation of adjusted scoring average is 1.16 strokes.[29] Therefore, these results are quite precise. Furthermore, differences in productivity

[29]This statistic is computed for all golfers who play in a least five Web.com Tour or PGA TOUR events. Skewness is 0.33 and kurtosis is 3.92. As skewness is above zero, adjusted scoring average is positive skewed (toward poor scores). As kurtosis is above 3, the distribution has fatter tails than a normal distribution.

are indistinguishable from zero for almost the entire period of analysis.[30]

In comparison to more typical labor market data, golf earnings represent a more direct measure of productivity since it is clear that they are based on performance. However, differential access to golf tournaments may lead to earnings differences conditional on performance. The PGA TOUR provides about ten times more prize money than the Web.com Tour. However, the competition is also stronger. These factors represent counteracting forces in terms of a golfer's expected earnings. If for the average qualifier, the effect of increased prize money outweighs the effect of stronger competition, then I expect treated golfers to earn more than control golfers even while holding quality of play constant. Given a positive effect on earnings and no effect on performance, the results suggest that there are rents to be gained from access to the PGA TOUR and that the initial earnings effects are the result of differential access to rents rather than differential quality of play.

In summary, Figures 2.8 and 2.9 provide strong evidence that both treatments have no significant long-term effects on earnings and do not affect productivity. Furthermore, Appendix Figures 2.24 and 2.25 show that these results hold for both younger and older golfers. These three pieces of evidence all support a rent-sharing as opposed to a human capital mechanism.

2.6.2 Treatment Effects on Employment

Figure 2.10 displays the estimated treatment effects on future PGA TOUR events played. For both experiments, we see large treatments effects in terms of PGA TOUR events played in the first year. As expected, the Q School treatment is initially stronger than the Web.com Tour ML treatment.

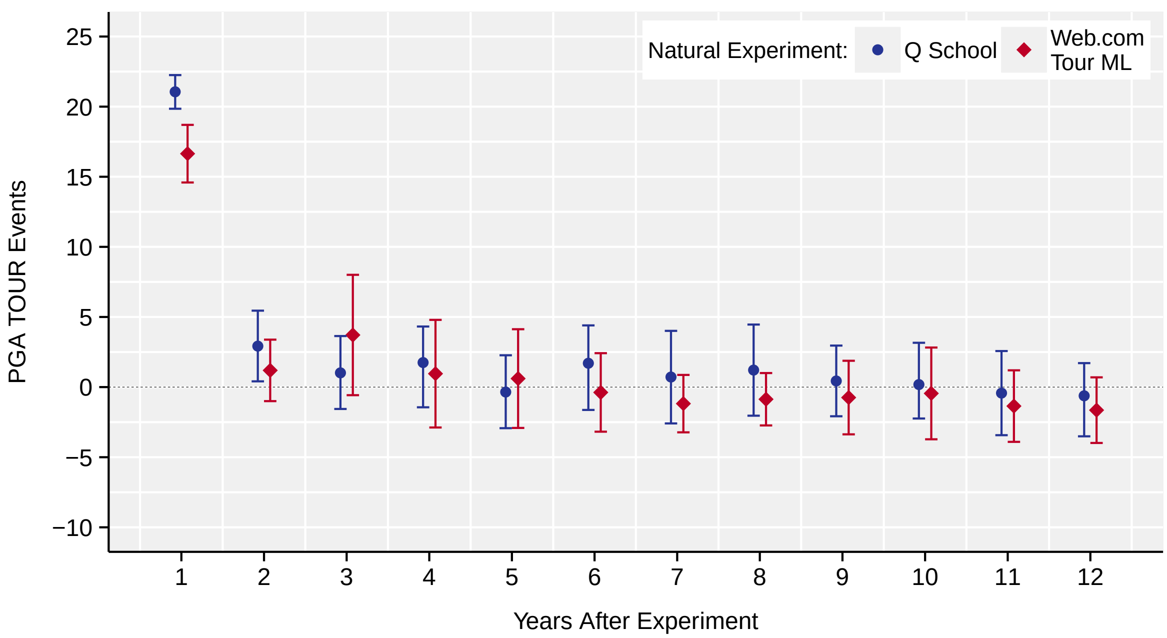
[30]There is an estimated statistically significant *decrease* in productivity for the Web.com Tour ML sample in the twelfth year after treatment. However, this effect is not robust across specifications. For instance, the result does not hold using the conventional RD estimates (i.e. no bias correction, Figure 2.37) nor does it appear when I apply local randomization methods to the Web.com Tour ML experiment(Figure 2.36). Also, it seems implausible that productivity effects would show up so long after treatment with no prior evidence of significant differences. For these reasons, and since there is large variance in the Web.com Tour ML at long time horizons, I conclude that there is not enough evidence to interpret this estimate as a robust, real effect.

Q School qualifiers are estimated to play in 21.1 more events than non-qualifiers, while Web.com Tour ML qualifiers play in 16.7 more events. These results are consistent with the findings of Section 2.6.1 where large earnings effects are found in the first year after treatment.

Figure 2.10 also shows that treatment effects quickly dissipate for both treatments. For the Q School treatment, there is evidence of a small effect in year two–2.9 events with a p-value of 0.055. However, for the remaining years the difference in the number of PGA TOUR events played is statistically insignificant. For the Web.com Tour ML treatment, the estimated effect is close to zero in year two but jumps up to 3.7 events in year three. The large variance around this estimate, however, renders it statistically insignificant at the 5% level with a p-value of 0.090. For the remainder of the observation period, estimated treatment effects on events played are statistically insignificant. Section 2.6.1 describes how the Web.com Tour ML treatment may be slightly more persistent in terms of earnings, whereas the Q School treatment appears to be slightly more persistent in terms of events played. There is no clear rationalization of these findings. My preferred interpretation is that there is some evidence for short-term and small persistent effects in both treatments, however, these should be considered transitory effects when compared to the shocks in the broader US labor market.

Figure 2.11 displays the estimated treatment effects on only the extensive margin of employment–the future probability of realizing any positive world earnings. For the Web.com Tour ML experiment, there appear to be essentially no effects on employment. In the first three years, I estimate very precise zero effects. Starting in the fourth year there is some variation in the estimated treatment effects across years, but there does not appear to be any systematic patterns and all treatment effects are statistically insignificant from zero at the 5% level, except for treatment effects estimated in years 11 and 12. Although not pictured, the estimated treatment effects jump back up to zero for years 13 through 16 (Appendix Table 2.7). Figure 2.12 plots the RD limit of the probability of future world employment on each side of the treatment threshold. In the Web.com Tour ML

Figure 2.10: Treatment Effects on Future PGA TOUR Events



Notes: *See notes to Figure 2.8.*

panel there is a sharp drop in employment in years 11 and 12 which abruptly returns to trend in year 13. If treated golfers retired early in year 11, then employment should be sustained at the low level. Given that years 11 and 12 are a temporary oscillation from trend, the estimated employment effects appear to be the result of an idiosyncratic factor such as injuries.[31]

In contrast to the Web.com Tour ML treatment, I estimate some significant employment effects for later years in the Q School experiment. Treated Q School golfers are about 20% more likely to be playing golf six to nine years after treatment. These effects are all significant at the 5% level. Furthermore, these results hold across age groups (Appendix Figure 2.29 and Appendix Table 2.12). Older golfers are more likely to play on the PGA TOUR (Appendix Figure 2.30) whereas the younger golfers are more likely to play on the Web.com Tour (Appendix Figure 2.31). There is also evidence that treated golfers play in more total US events (PGA TOUR + Web.com Tour) in years six through eight (Appendix Figure 2.20 and Appendix Table 2.12). The estimates on total US events during these years are not as significant, however, suggesting that the golfers who extend their careers play in few events.

Given these employment effects, a natural question is whether treated golfers play in more events or non-treated golfers play in less. For some indication I look to the employment levels on each side of the treatment threshold plotted in Figure 2.12. First, note that in the Web.com Tour ML panel the decline in employment appears approximately linear. If we expect the same type of linear trend through the first 9 years after the Q School treatment, then it is the treated golfers that seem to deviate. In other words, the treated golfers appear to playing longer than expected.

The total effect of treatment includes both earnings and employment effects. Estimation of the total effect requires a specification in levels rather than log as log earnings are undefined at zero. This specification is likely to add noise to estimated ATE due to the skewed nature of the earnings

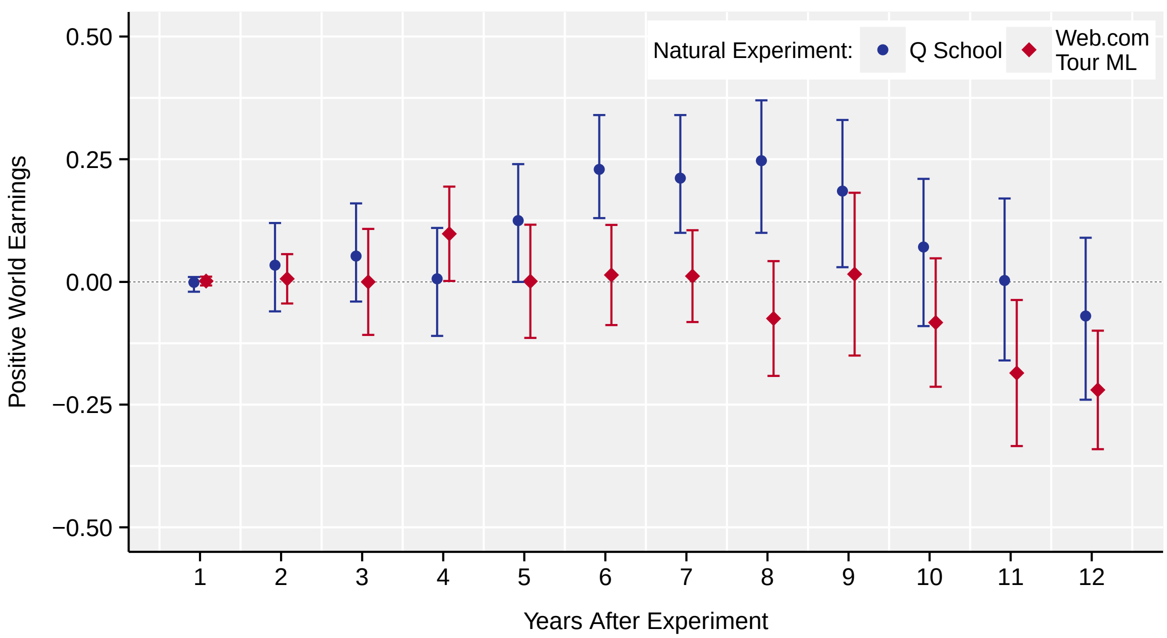
[31]Futhermore, if I use conventional RD local linear methods (i.e. if I don't apply the bias correction of Calonico et al. (2014)), the signifcance of the treatment effects estimated in years 11 and 12 for the Web.com Tour ML disappears.

distribution as discussed in Section 2.3.2. Appendix Tables 2.9 and 2.14 show the results of the total effect of treatment for the Web.com ML and Q School experiments, respectively. Among years six through nine, the *World Earnings* panel of Appendix Table 2.14 reports a significant total earnings effect only in year six (p-value=0.039) and a moderately significant effect in year eight (p-value=0.103). Even though some treated Q School golfers have longer careers, these golfers do not appear to accumulate enough earnings to produce strong evidence of a positive total earnings effect.

A question remains as to how the treated Q School golfers stay employed given a lack of earnings. Perhaps treated golfers feel that they are PGA TOUR quality golfers and, as a reulst, are reluctant to quit.[32] Another explanation may be that Q School is itself an entertainment spectacle where golfers gain notoriety which increases their chances of obtaining sponsor's exemptions into future events. The fact that a book about Q School (Feinstein (2007)) became a national bestseller lends some support to this hypotheses. However, these hypotheses remain speculative without greater empirical support.

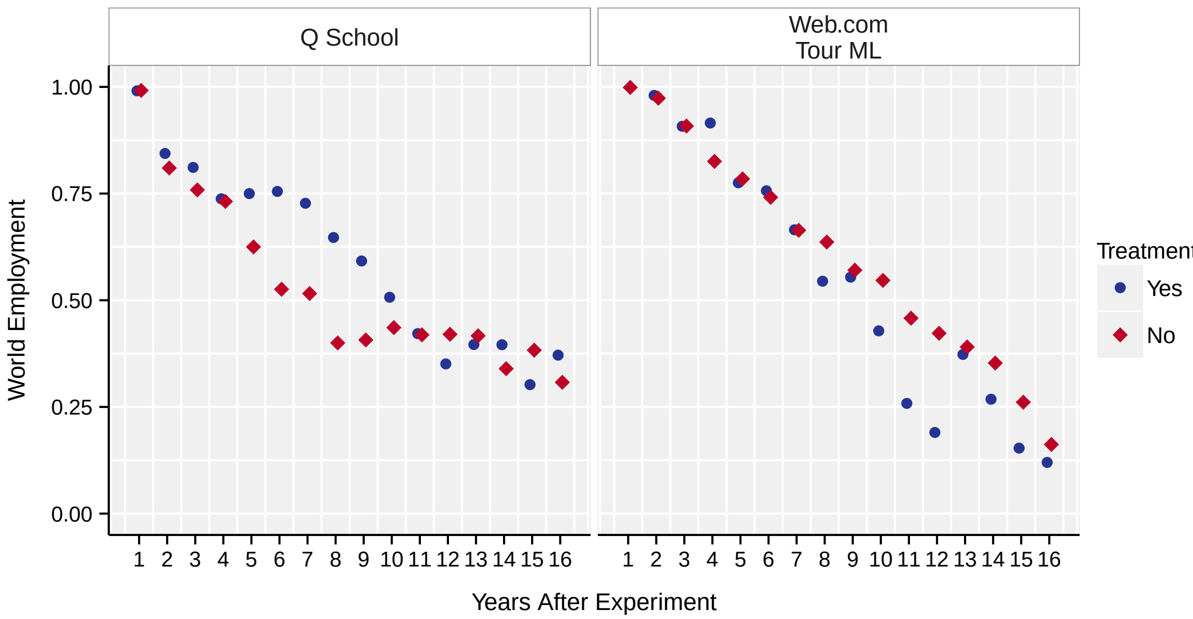
[32]Of course, why this would apply to Q School golfers and not Web.com Tour ML golfers is unclear.

Figure 2.11: Treatment Effects on Future Probability of Positive World Earnings



Notes: *See notes to Figure 2.8.*

Figure 2.12: Estimated Threshold Limits for Future Probability of Positive World Earnings



Notes: *See notes to Figure 2.8.*

2.7 Comparison to US Labor Market

Employment shocks in the broader US labor market have been shown to have persistent earnings effects, such as job displacements (von Wachter et al., 2009) and graduating during a recession (Oreopoulos et al., 2012), whereas in the golf setting, earnings effects are temporary despite a large initial shock. How then can we explain the differences in these results? I provide two pieces of evidence to compare the professional golf labor market to the US labor market. First, I discuss differences in transition rates between employment states. Next, I compare earnings shocks in the golf versus the US labor market through decompositions of total variance into transitory and permanent components.

2.7.1 State Transitions

In order to understand the mechanisms behind the treatment effects, annual transitions between employment states are informative. For instance, suppose a PGA TOUR member is defined as any golfer who plays in at least 20 official PGA TOUR events in a year. Let t denote the year of treatment. For the Web.com Tour ML experiment, the RD limit estimate for PGA TOUR membership in year $t+1$ is 90% for treated versus 28% for control. Of the treated players that earned tour membership in year $t+1$, the RD limit estimate of the percentage that retain their membership in $t+2$ is 32%. Of the control golfers who did not earn tour membership in year $t+1$, the RD limit estimate of the percentage that earn membership in year $t+2$ is 27%.[33] Therefore, for a population of golfers with comparable ability, the exit rate off of the PGA TOUR players is estimated as 0.68 and the entry rate to the PGA TOUR is 0.27.

Using administrative US Social Security Administration earnings data, Smith et al. (2017)

[33] For the Q School treatment, 82% of treated golfers and just 3% of control golfers play in at least 20 PGA TOUR events. Of the 82% who are members in year $t+1$, 37% retain membership in year $t+2$. Of the 97% of control players who fail to earn membership in year $t+1$, 24% advance to the PGA TOUR in year $t+2$.

report annual job transition rates. For a males aged 35, they report an average nonemployment to employment transition rate of 0.2, an employment to nonemployment transition rate of 0.075, and a job-to-job transition rate of 0.2. Furthermore, they compute the probabilities of moving up and down the job ladder using a measure of firm quality.[34] Conditional on switching firms, they report a probability of moving up the job ladder of 0.64 and a probability of moving down the job ladder of 0.36.

Putting these numbers together I estimate the probability of transitioning from a good to a bad state as: 0.075 (E→N) + 0.2*0.36 (down the job ladder) = 0.15. The probability of transitioning from a bad to a good state is then: 0.2 (N→E) + 0.2*0.64 (up the job ladder) = 0.33.[35] The transition rates from bad to good states are similar: 0.27 in golf versus 0.33 in the US labor market. However, when in terms of exit rates, golfers have a transition rate of 0.68, whereas workers have a displacement rate of approximately 0.15. Golfers are about 4.5 times more likely to suffer a negative transition. Relative to a professional golfer, the average worker appears to have some insulation from negative shocks. This insulation may contribute to the persistence of the effects we see from temporary employment shocks.

2.7.2 Transitory versus Permanent Earnings Shocks

Another way to understand the golf earnings process in relation to the broader US labor market is to decompose earnings variance into transitory and permanent components. There is a large literature discussing optimal specifications to estimate error components models of earnings processes.[36] However, these models can be well approximated by reduced-form methods.

---

[34] They use firm fixed effects as originally laid forth in Abowd et al. (1999) as their measure of firm quality.

[35] These calculations implicitly assume that everybody in the nonemployment state wants to work. Of course this is a crude approximation. Relaxing this assumption will increase the N→E rate and decrease the E→N rate above.

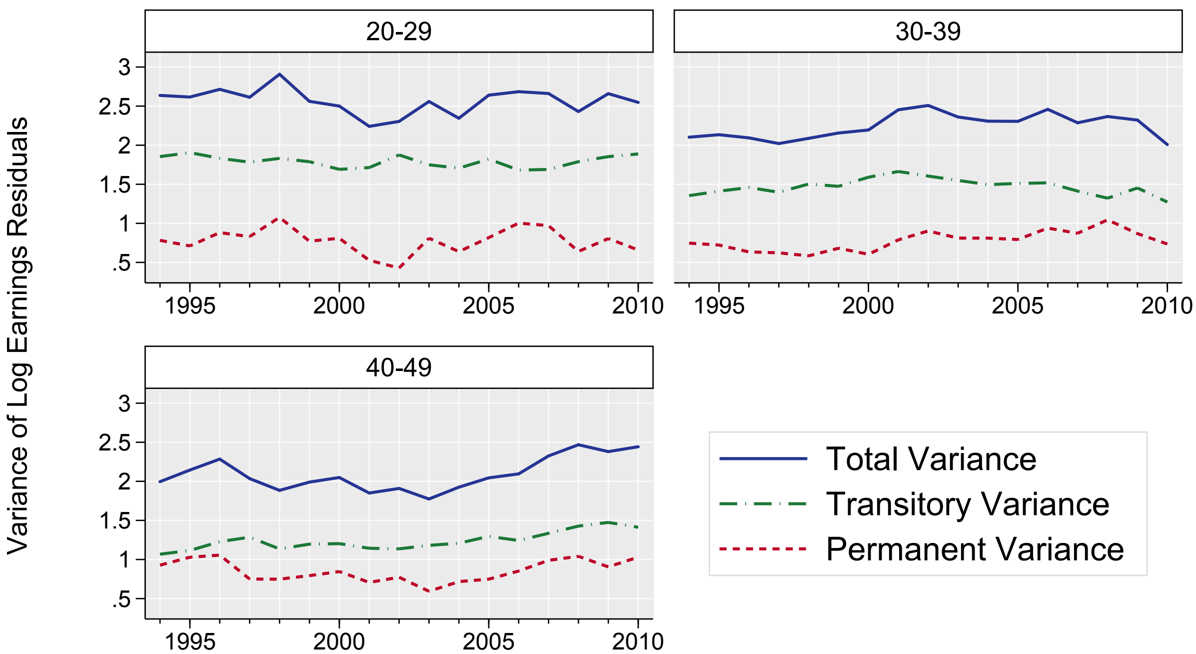
[36] See Moffitt and Gottschalk (2011) for one overview.

I apply the *window averaging method* of Gottschalk and Moffitt (1994) and Moffitt and Gottschalk (2011) to estimate the transitory and permanent variance components in golf earnings. The first step is to compute earning residuals based on a regression of log earnings on education, race, an age polynomial, and interactions between these variables. This regression is estimated separately for each calendar year. The idea behind using earnings residual instead of plain earnings is to highlight shocks to the earnings process rather than cross-sectional differences in average earnings levels or growth paths. Within a window of nine years, four years before and four years after, all residuals are averaged to obtain an estimate of the individual's permanent component. Nine years is a rough estimate of the amount of time it takes for a transitory shock to dissipate. To estimate the transitory component, they take the difference between the residual for each individual in each year and the individual's average residual. Then textbook formulas for the random effects model are used to compute the variances of the two components.

I adapt this method to the golf setting with the caveat that my residuals are based on a regression of log earnings on polynomials in lifetime average OWGR and age, interactions between these variables, and year effects. Lifetime average OWGR is the closest variable in this context to an education variable. Both are meant to proxy for underlying skills but they may be significantly different since lifetime average OWGR is an outcome and education is more of an input.

Figure 2.13 displays transitory and permanent residual variance components by age group and over time. Note that the earnings process seems quite stable over time with no strong trends in any age group for total, transitory, or permanent variance. This can be interpreted as evidence that the earnings process, including the level of competition and compensation structure, has remained stable over time. Also note that residual variance appears to decline with age. The more substantial finding, however, is that golfer earnings appear to have a larger proportion of transitory variance than earnings from the broader US labor market. Using the Michigan Panel Study on Income Dynamics, Moffitt and Gottschalk (2011) find about 35-40% of the total residual variance of 30-39

Figure 2.13: Permanent and Transitory Earnings Components: Window Averaging Method



Notes: *I apply the window averaging method of Moffitt and Gottschalk (2011). Earnings residuals are computed from a regression of earnings on polynomials in average lifetime OWGR and age, interactions between these variables, and year fixed effects. Permanent components are computed as the average of the residuals within each nine years window. Transitory components are the difference between the residual and the permanent component. Golfers must be in a nine year window for a minimum of four years.*

year olds from 1980 to 2000 is transitory.[37] As another example, using a similar methodology with IRS earnings data DeBacker et al. (2013) find that transitory variance accounts for about 17% of total residual variance for 25-60 year old males from 1989 to 2007.[38] In contrast, for golfers between ages 30 and 39 I find that transitory variance comprises approximately 65% of the total residual variance over the full period. These results provide an indication that golfers may be able to recover more quickly from shocks that the average worker. However, this evidence is only suggestive since there are no comparable regressions to produce the residual variance.

Economists often interpret the transitory component of earnings variance as a reflection of market competition. For example, Moffitt and Gottschalk (2011) state that an increase in transitory variance could be caused by "an increase in product or labor market competitiveness, a decline in regulation and administered prices, a decline in union strength, increases in temporary work or contracting-out or self-employment, and similar factors." Professional golf is an example of a labor market that suits this description. Profession golf is highly competitive and earnings are directly tied to compensation. Golfers are not insulated from negative shocks and, therefore, face a high risk of job displacement.

2.8 Conclusion

I study two natural experiments in professional golf which provide exogenous shocks to golfers' employment states. Initially, I find large effects on both earnings and employment, with treated golfers benefiting from a 94% and 64% earnings increase and playing in 21.1 and 16.7 more PGA TOUR events in the Q School and Web.com Tour ML experiments, respectively. However,

[37]Moffitt and Gottschalk (2011) do not present their results in table form, but only plot the results. Therefore, I imputed these number using Figures 5 and 8 in years 1980, 1990, and 2000 from their study.

[38]Once again, these numbers are not explicitly provided in table form but are inferred from Figure 3b of DeBacker et al. (2013).

in each experiment these effect quickly dissipate. There is no evidence of long-term effects on earnings in either experiment. There are some long-term employment effects from the Q School experiment, but these golfers are not productive enough to create any significant total earnings differences. Furthermore, I find that there are no productivity differences across treatment states. Therefore, I attribute all benefits of the treatment to the direct effects of increased access rather than improved performance. These results favor a rent-sharing mechanism rather than a human capital mechanism.

In comparison to some temporary employment shocks studied in the broader labor market, namely job displacement and graduation during recessions, these effect are less persistent. To establish a comparison between professional golf and a general labor market, I document transition rates between employment states. I find that golfers have a much higher exit rate from good states than found in the US labor market. These high exit rates may be the result of lower job market frictions, higher competition, or less insulation from negative shocks. Further, I document that a greater proportion of the variance in earnings shocks is transitory than is found in the broader labor market. This is yet another indication of a more competitive or mobile environment.

Increased mobility or a lack of insulation from negative shocks, may then explain the lack of persistence from employment shocks in the golf setting. Relative to golfers, the average worker appears more insulated from negative shocks. In general, this may represent an improvement in worker welfare, but will result in persistent costs from job displacements given that it takes time to find a comparable job.

This observation has implications for the wider labor market. For instance, Davis and von Wachter (2011) find that the negative effects of job displacement are greater in recessions. A standard explanation for the cyclical nature of these effects rests on the observation that job offer rates decline during recessions. However, recessions are often short lived. A more lasting effect may be due to the fact that those displaced in recessions come, on average, from more stable jobs.

Therefore, their counterparts are less likely to become displaced in the future. Furthermore, we may expect workers from countries that offer greater protections from firing to experience greater costs of job displacement, conditional on job finding rates.

2.9 Appendix: Additional Figures

Figure 2.14: Average Official Tournament Purse: PGA TOUR and Web.com Tour

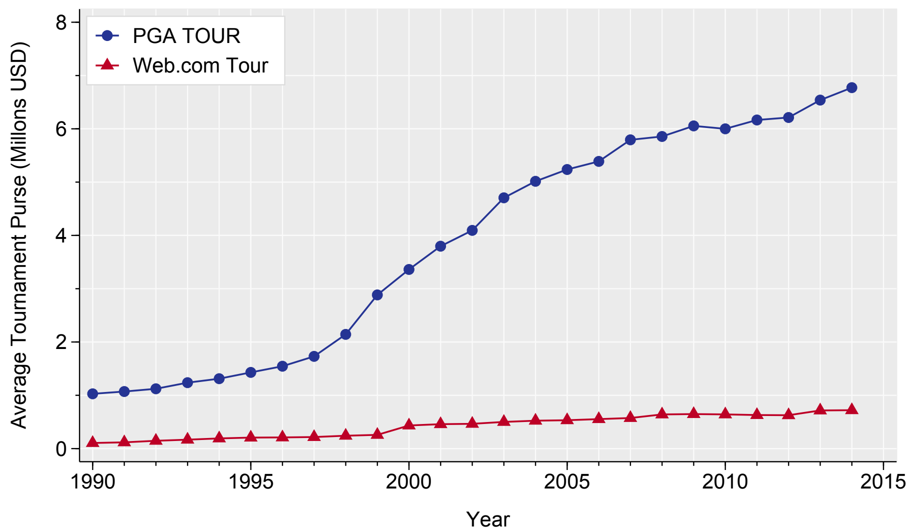


Figure 2.15: Ratio of Web.com Tour to PGA TOUR Average Prize Money

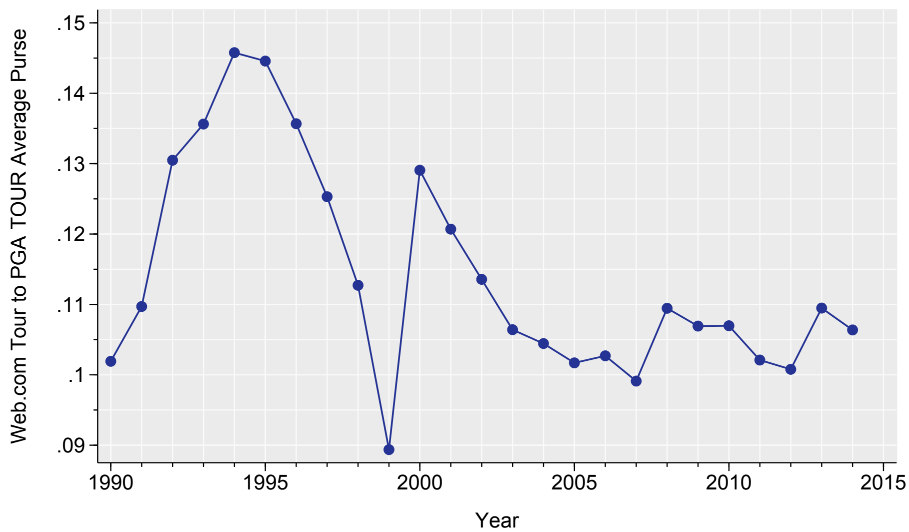


Figure 2.16: PGA TOUR Cards Awarded by Experiment

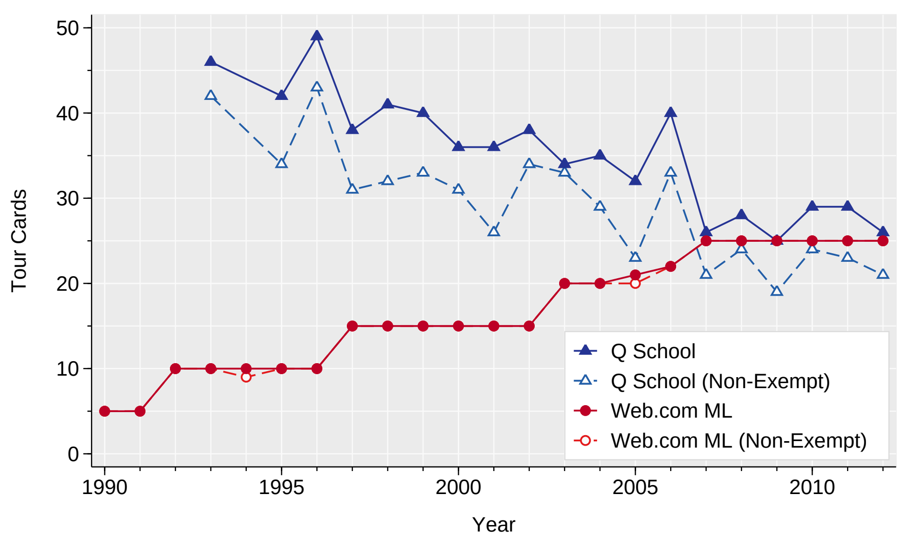
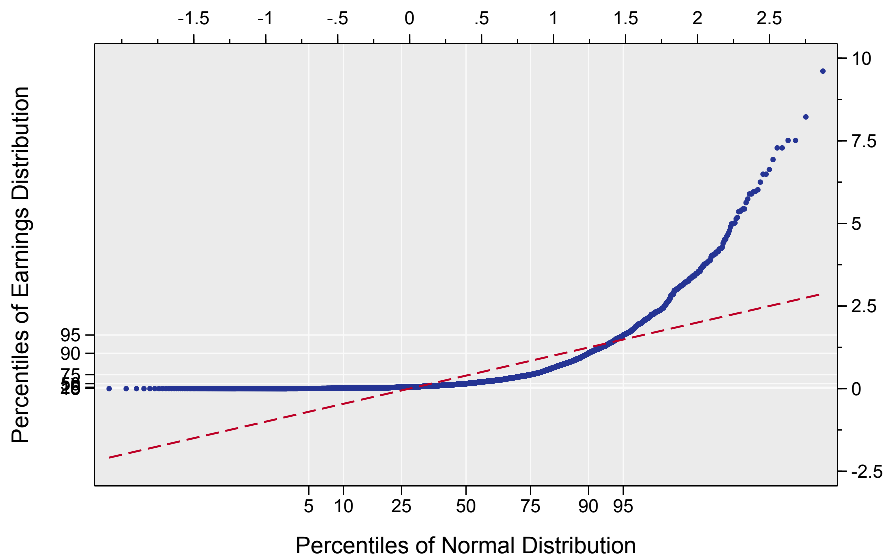
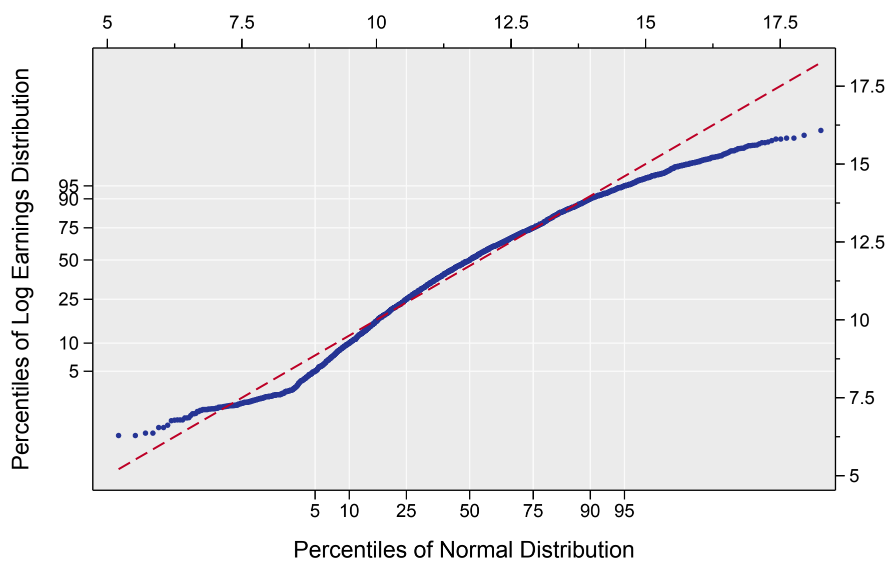


Figure 2.17: Earnings Distributions Relative to a Normal Distribution for Web.com Tour ML Experiment

(a) Future Earnings Distribution

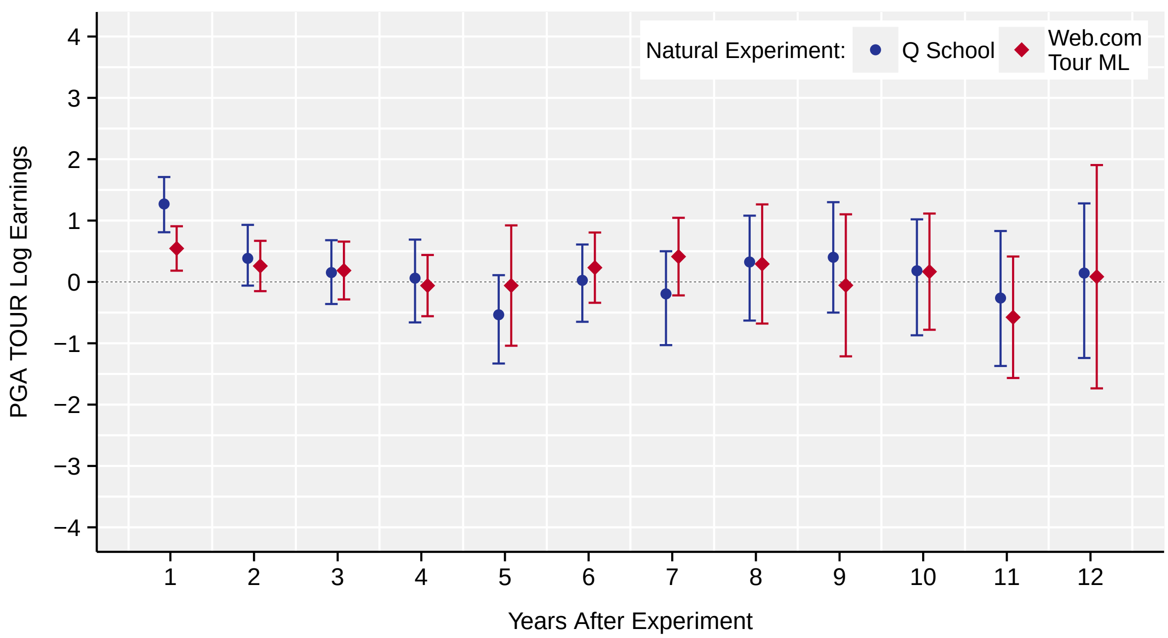


(b) Future Log Earnings Distribution



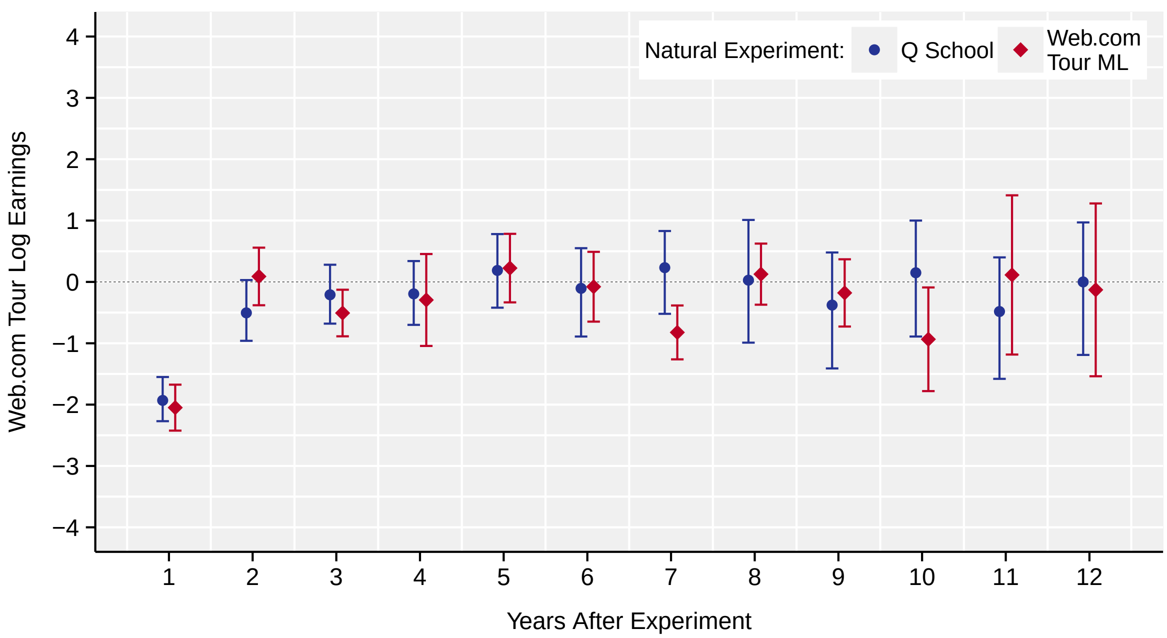
Notes: *Earnings from 1, 2, 3, 4, 5, 10, 15, and 20 years after treatment are aggregated and plotted. The left and bottom axes show the location of the notes percentiles of the distribution. The top and right axes show the scales. Earnings are in millions of dollars.*

Figure 2.18: Treatment Effects on Future PGA TOUR Earnings



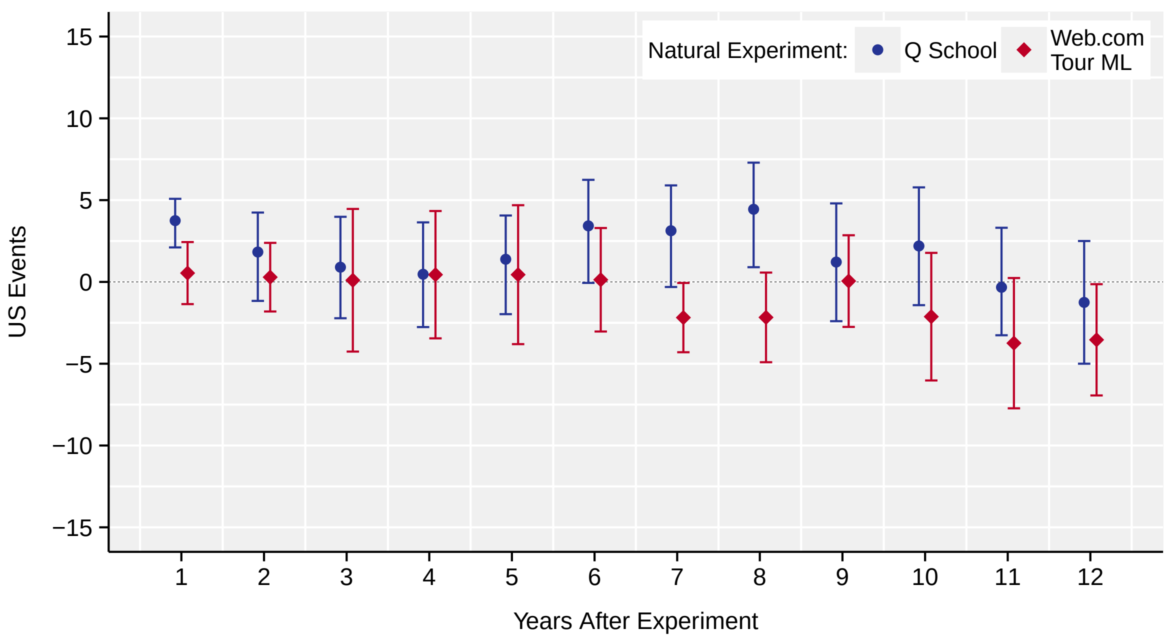
Notes: *See notes to Figure 2.8.*

Figure 2.19: Treatment Effects on Future Web.com Tour Earnings



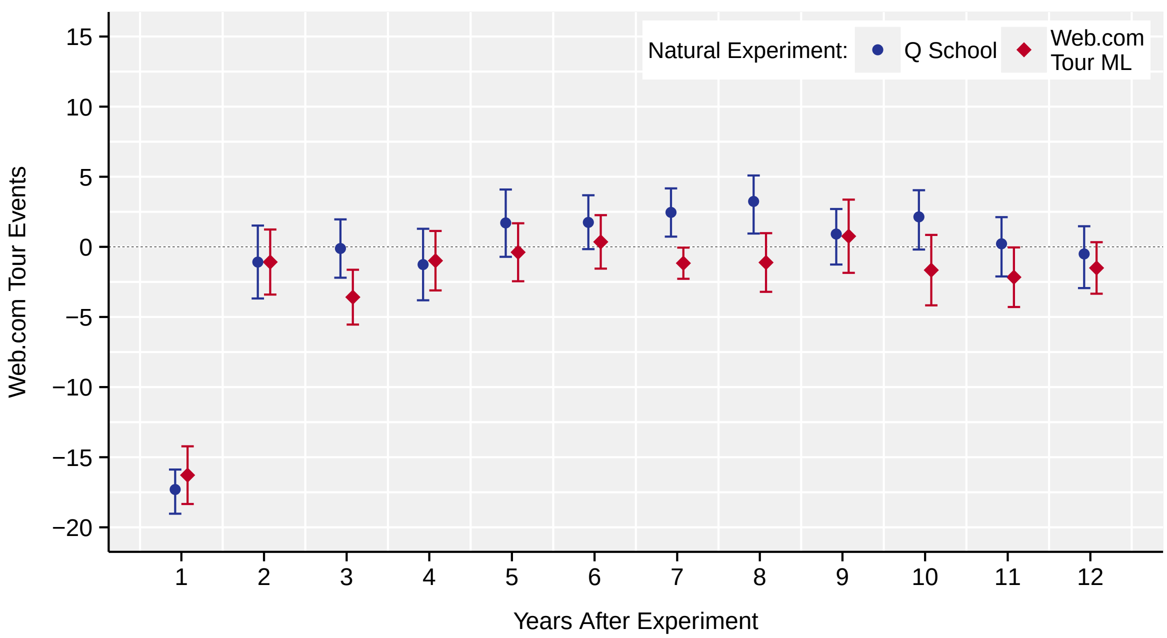
Notes: *See notes to Figure 2.8.*

Figure 2.20: Treatment Effects on Future US Events



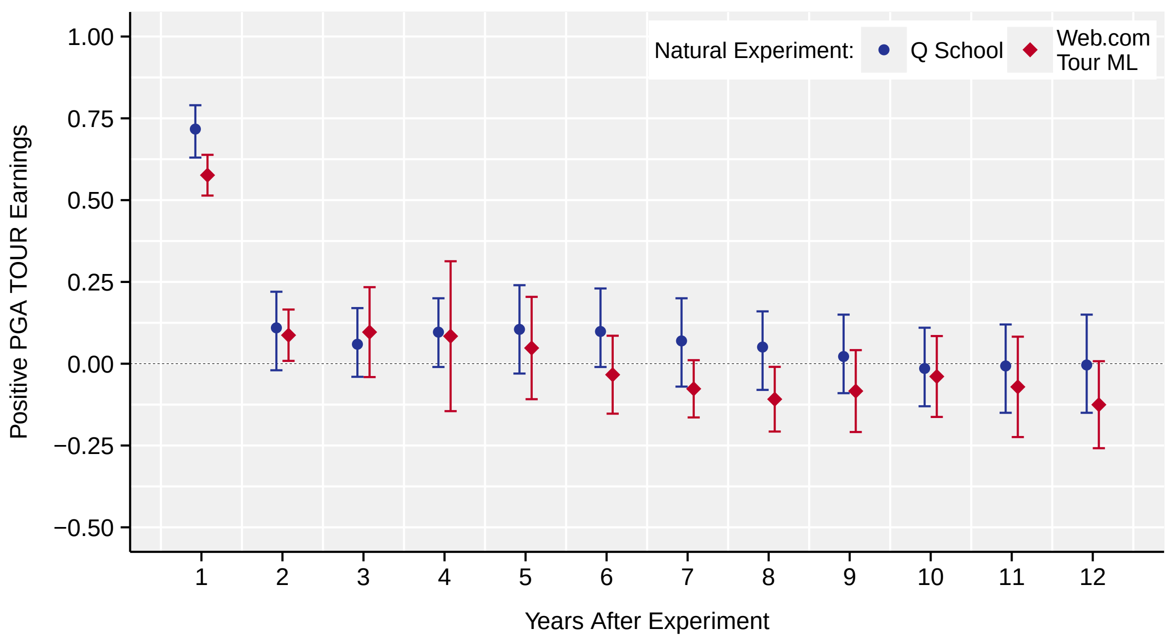
Notes: *See notes to Figure 2.8.*

Figure 2.21: Treatment Effects on Future Web.com Tour Events



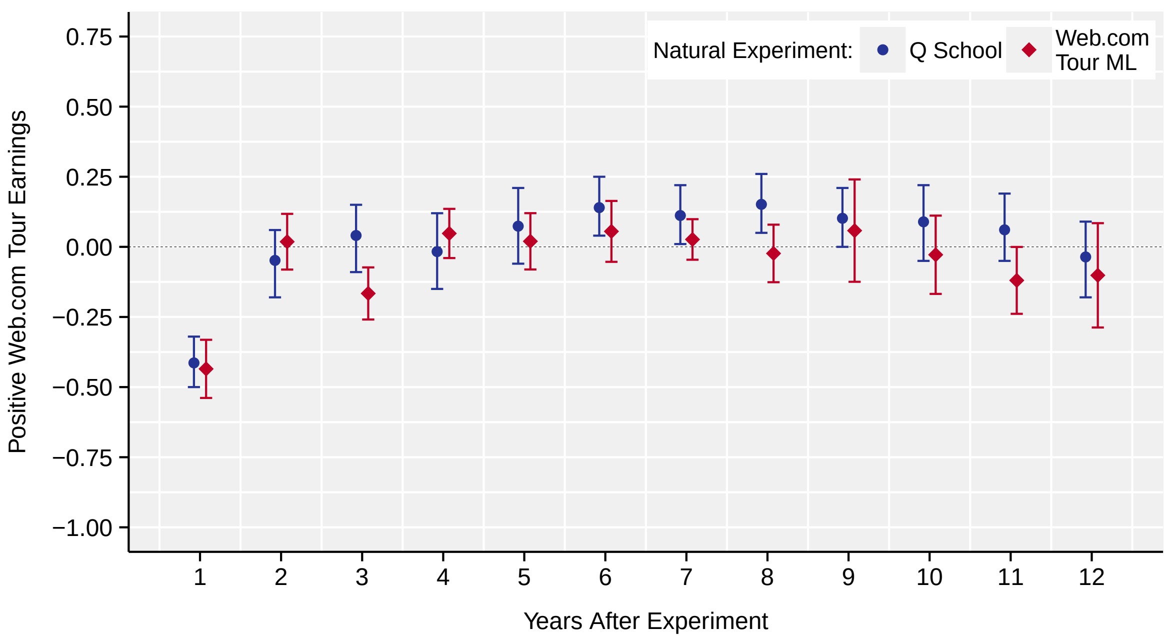
Notes: *See notes to Figure 2.8.*

Figure 2.22: Treatment Effects on Future Probability of Positive PGA TOUR Earnings



Notes: *See notes to Figure 2.8.*

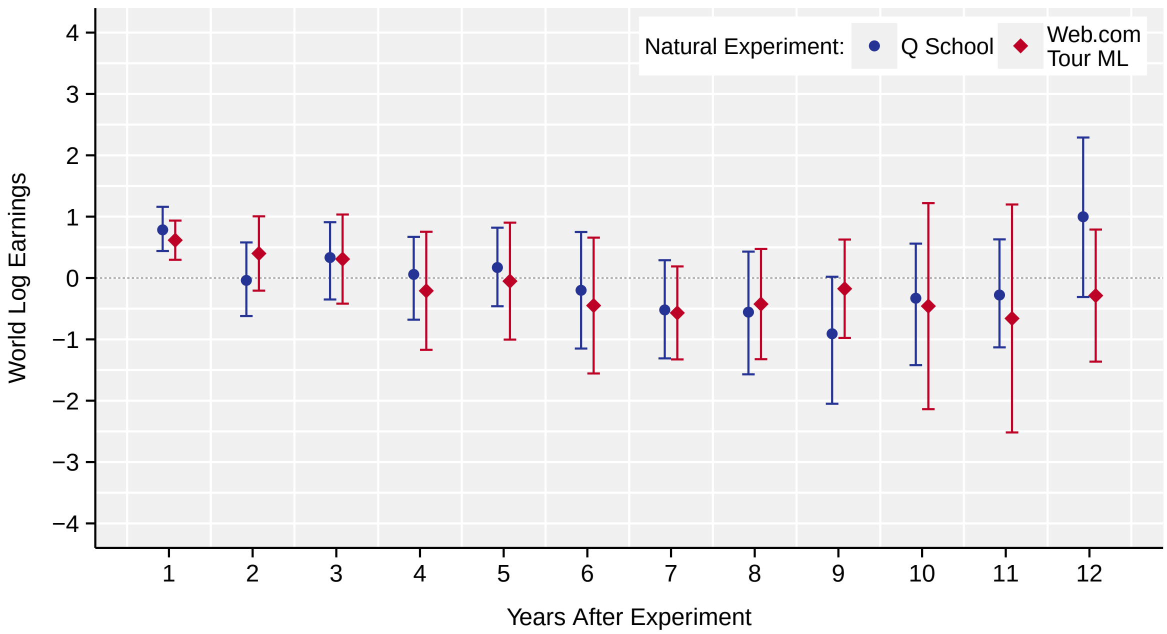
Figure 2.23: Treatment Effects on Future Probability of Positive Web.com Tour Earnings



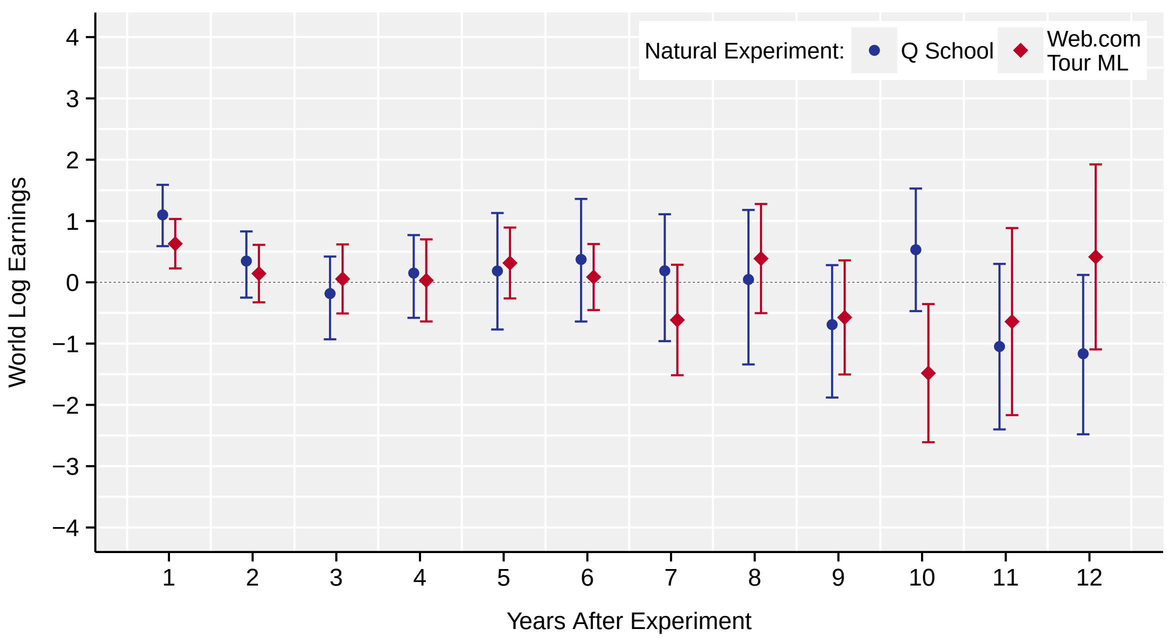
Notes: *See notes to Figure 2.8.*

Figure 2.24: Treatment Effects on Future World Earnings by Age

(a) Younger Golfers



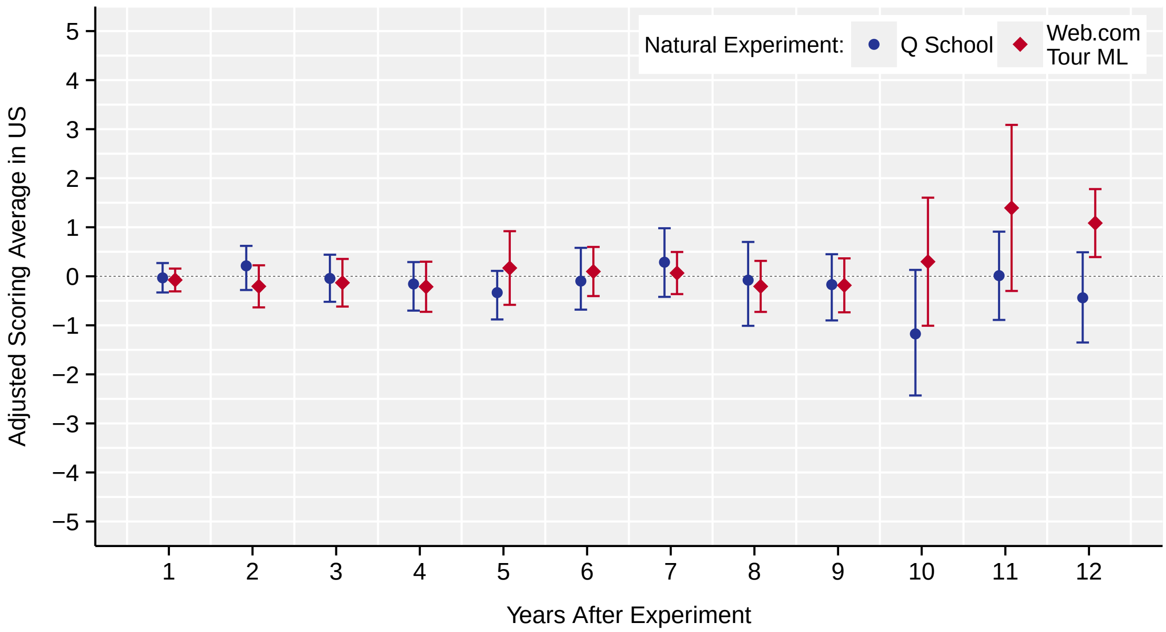
(b) Older Golfers



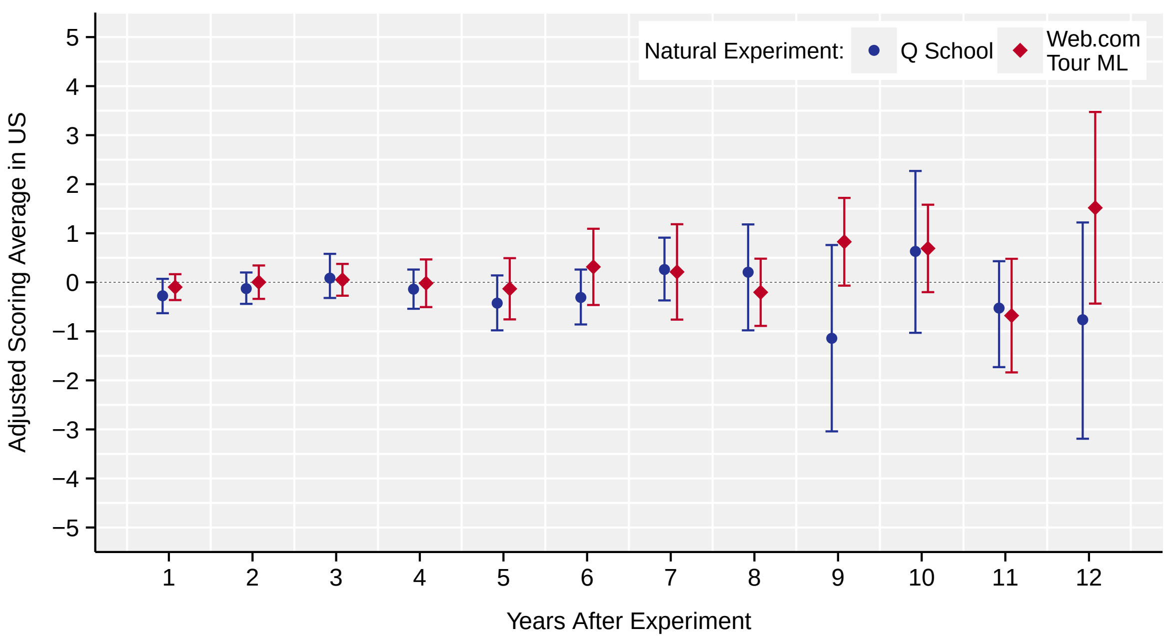
Notes: *See notes to Figure 2.8. Younger golfers are less than or equal to 30 years.*

Figure 2.25: Treatment Effects on Future Scoring Average in the US by Age

(a) Younger Golfers



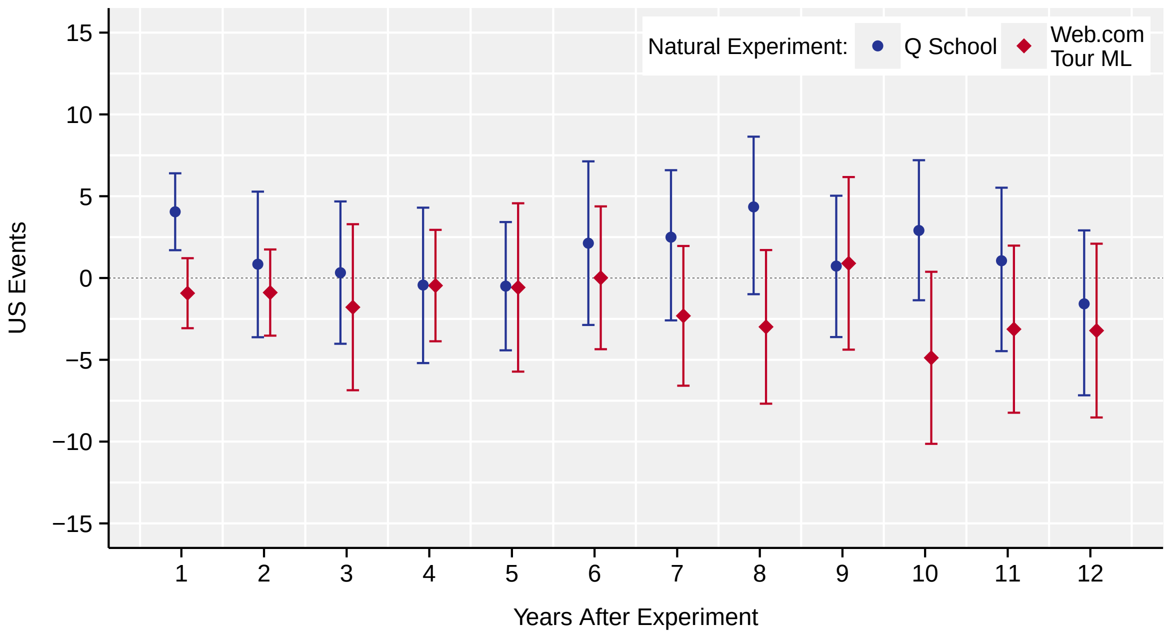
(b) Older Golfers



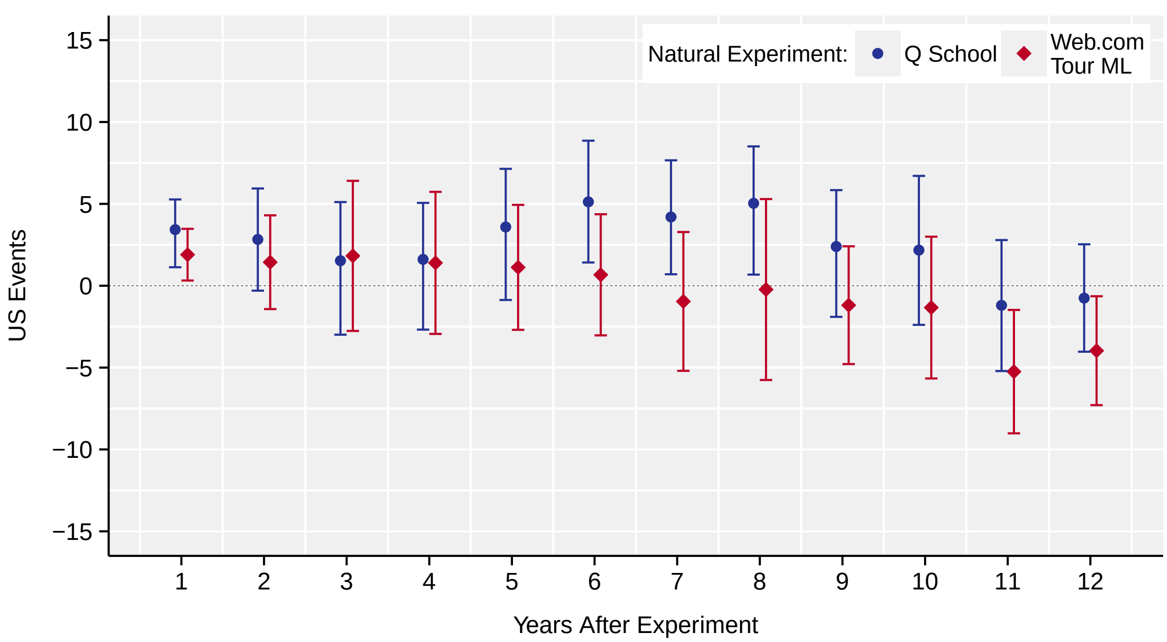
Notes: *See notes to Figure 2.8. Younger golfers are less than or equal to 30 years.*

Figure 2.26: Treatment Effects on Future US Events by Age

(a) Younger Golfers



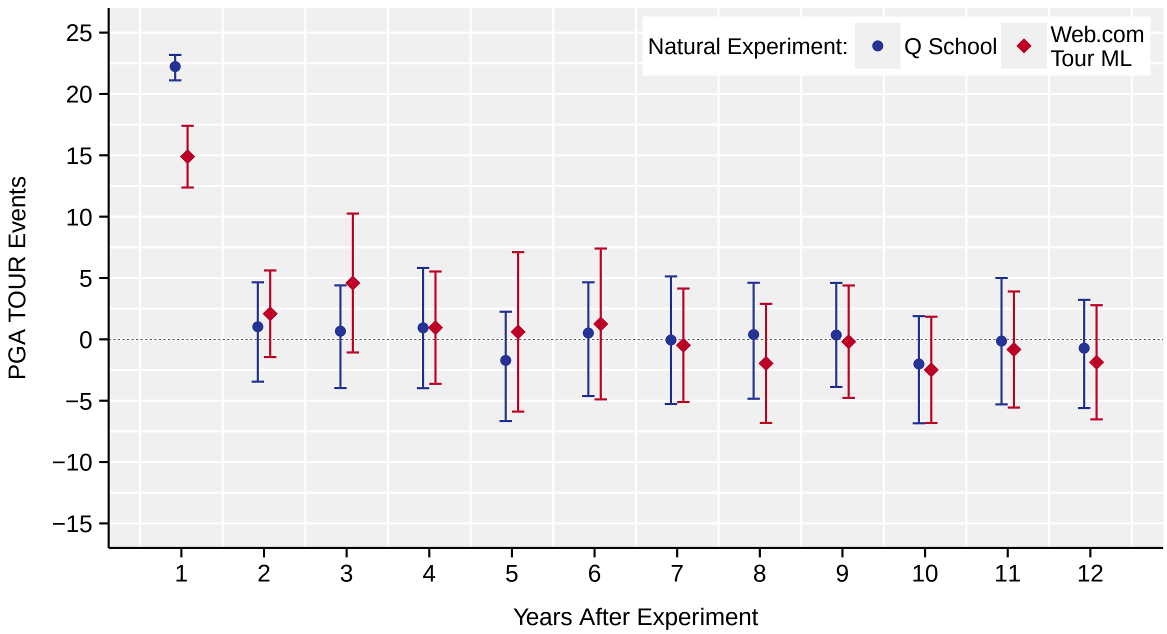
(b) Older Golfers



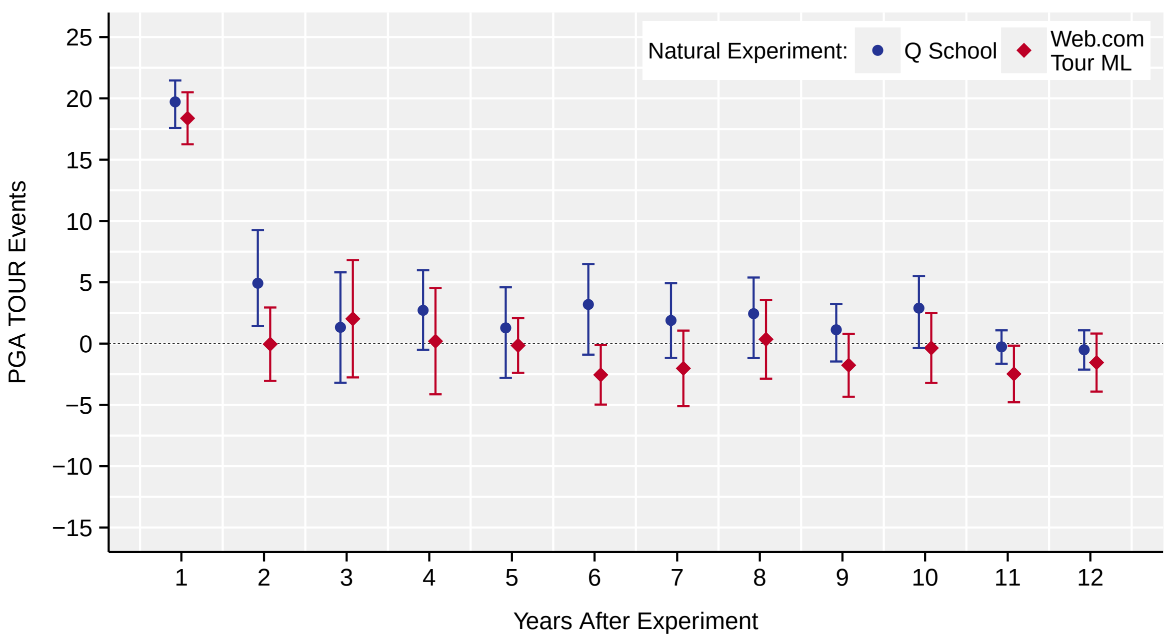
Notes: *See notes to Figure 2.8. Younger golfers are less than or equal to 30 years.*

Figure 2.27: Treatment Effects on Future PGA TOUR Events by Age

(a) Younger Golfers



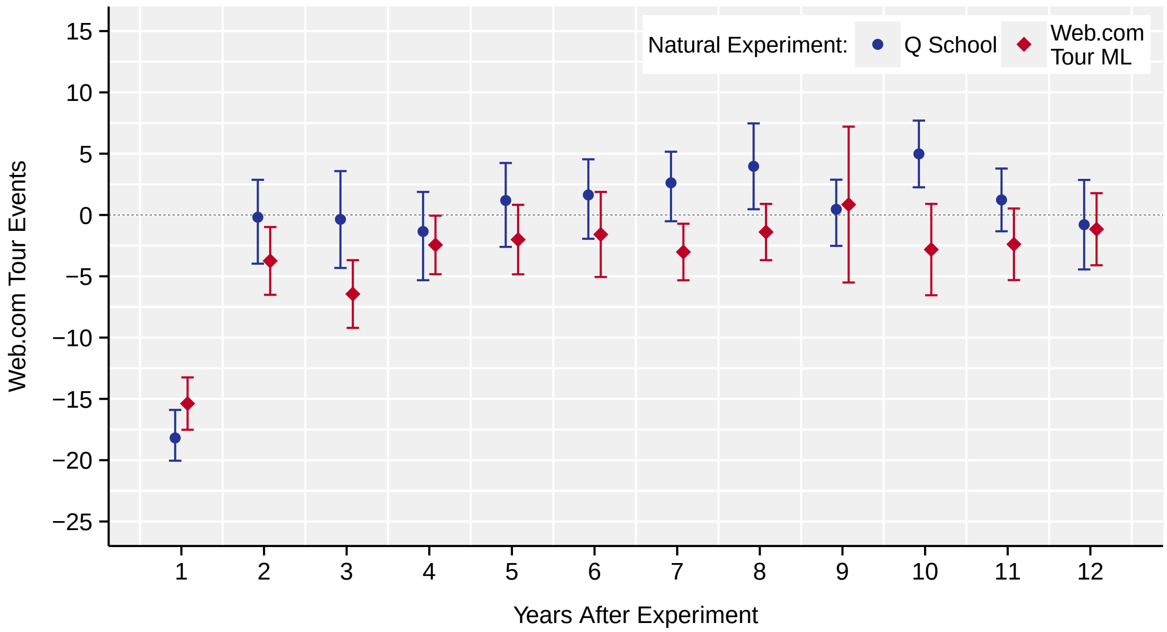
(b) Older Golfers



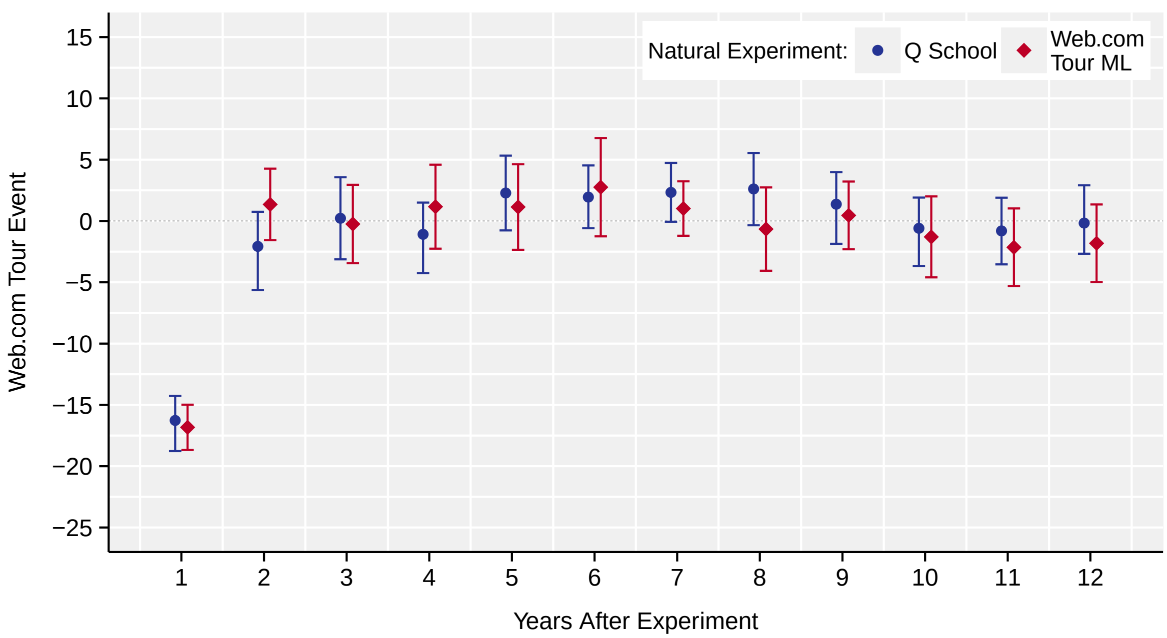
Notes: *See notes to Figure 2.8. Younger golfers are less than or equal to 30 years.*

Figure 2.28: Treatment Effects on Future Web.com Tour Events by Age

(a) Younger Golfers



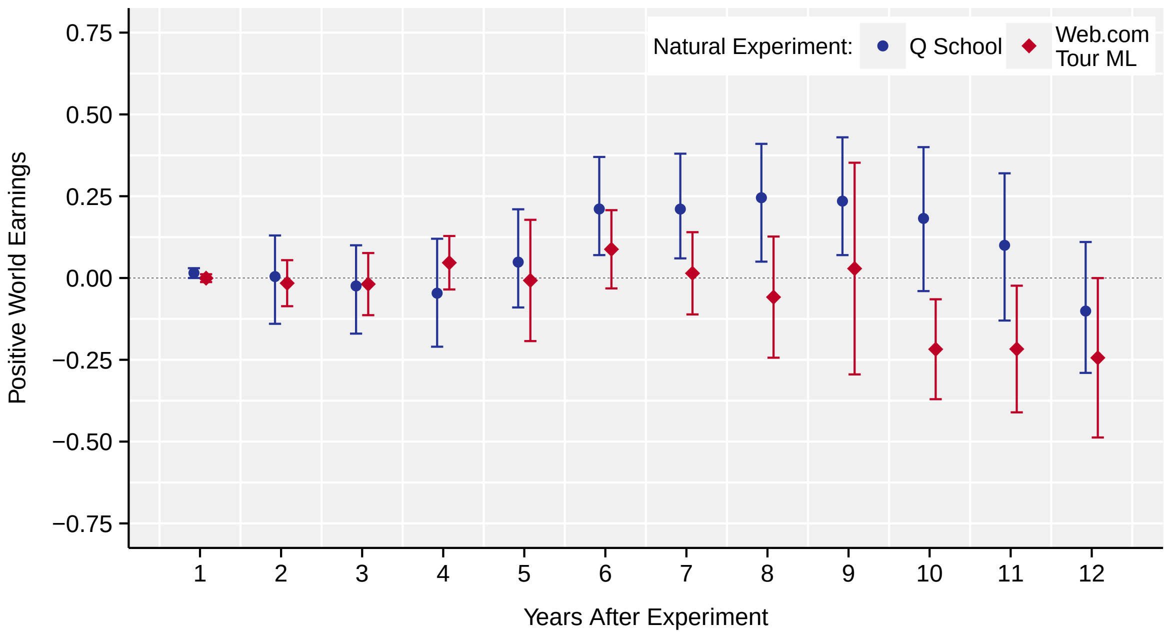
(b) Older Golfers



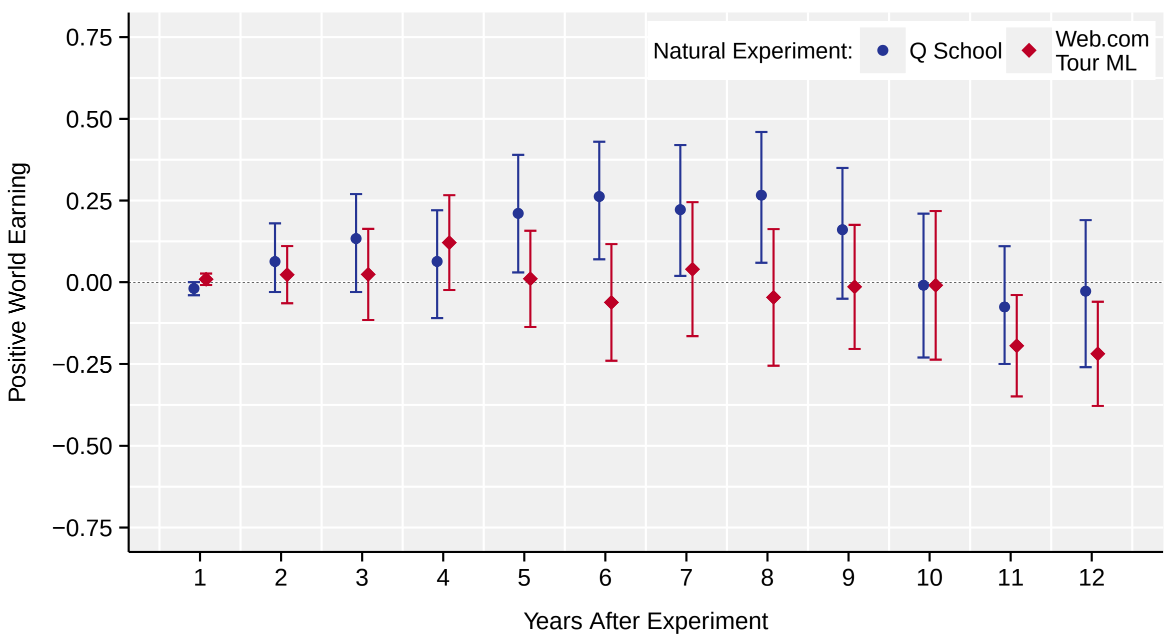
Notes: *See notes to Figure 2.8. Younger golfers are less than or equal to 30 years.*

Figure 2.29: Treatment Effects on Future Probability of Positive World Earnings by Age

(a) Younger Golfers



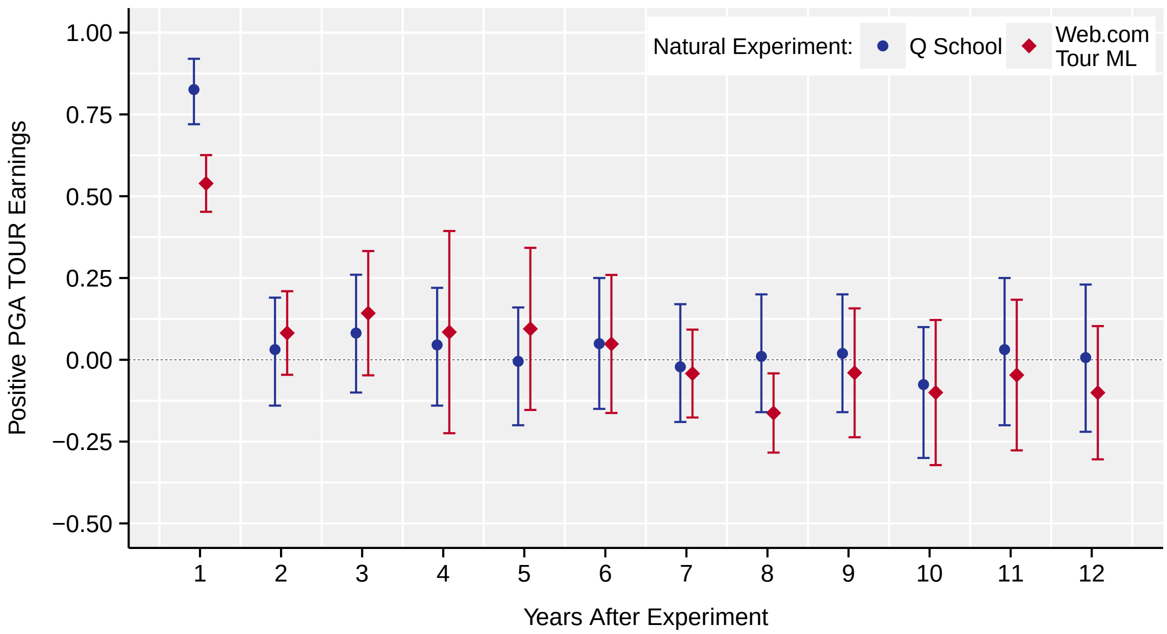
(b) Older Golfers



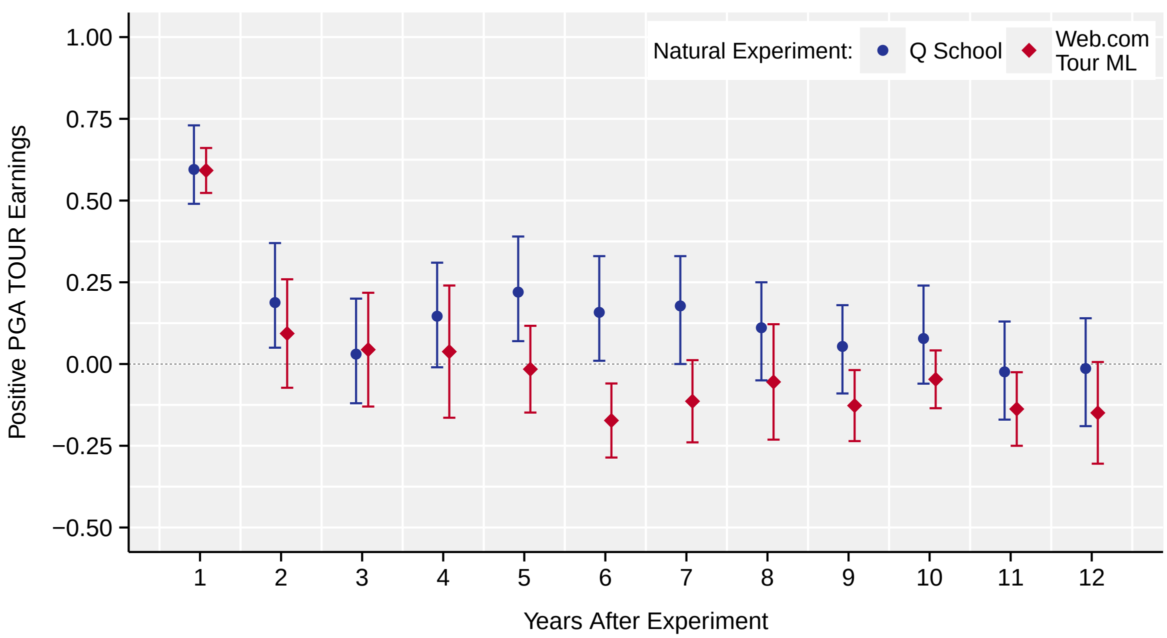
Notes: *See notes to Figure 2.8. Younger golfers are less than or equal to 30 years.*

Figure 2.30: Treatment Effects on Future Probability of Positive PGA TOUR by Age

(a) Younger Golfers



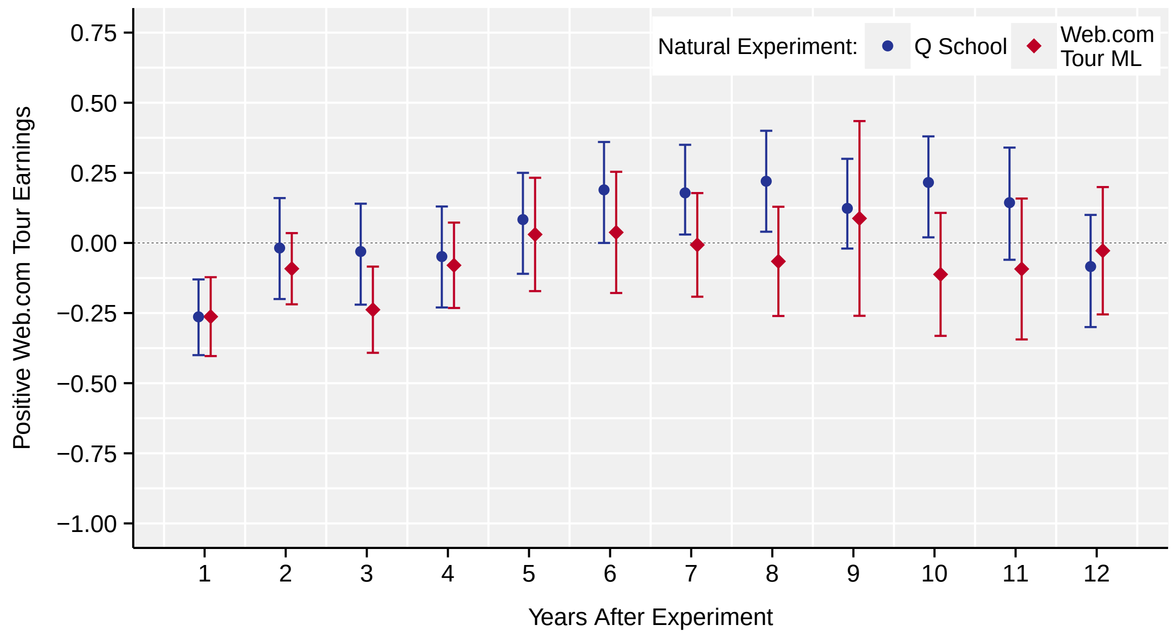
(b) Older Golfers



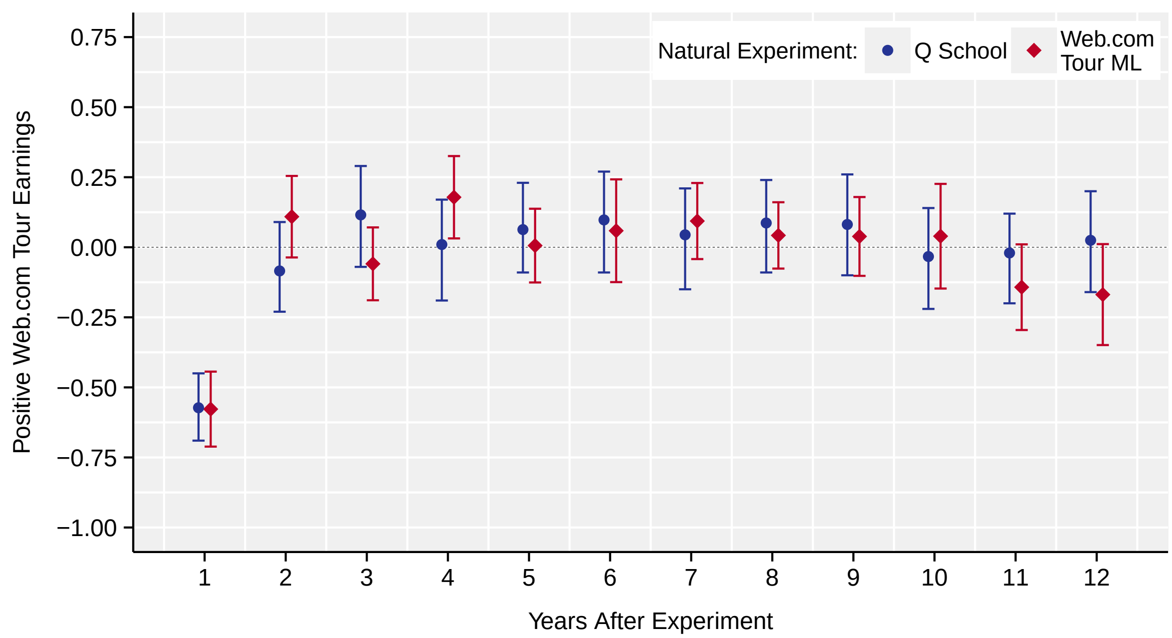
Notes: *See notes to Figure 2.8. Younger golfers are less than or equal to 30 years.*

Figure 2.31: Treatment Effects on Future Probability of Positive Web.com Tour by Age

(a) Younger Golfers



(b) Older Golfers



Notes: *See notes to Figure 2.8. Younger golfers are less than or equal to 30 years.*

2.10 Appendix: Additional Tables

Table 2.3: Available Worldwide Earnings Data

Tour	Years Available
PGA TOUR	1980-2014
European Tour	1980-2014
Web.com Tour	1990-2014
Challenge Tour	1990-2014
Japan Golf Tour	1985-2014
PGA Tour of Australasia	1980-2014
Asian Tour	1995-2014
Sunshine Tour	1991-2014

Table 2.4: Time and Age Composition of Estimates by Duration Post Treatment

Years After Treatment	Experiment Estimation Years		Potential Estimation Ages
	Q School	Web.com ML	
1	1993, 1995-2012	1990-2012	17-54
2	1993, 1995-2012	1990-2012	17-53
3	1993, 1995-2011	1990-2011	17-52
4	1993, 1995-2010	1990-2010	17-51
5	1993, 1995-2009	1990-2009	17-50
6	1993, 1995-2008	1990-2008	17-49
7	1993, 1995-2007	1990-2007	17-48
8	1993, 1995-2006	1990-2006	17-47
9	1993, 1995-2005	1990-2005	17-46
10	1993, 1995-2004	1990-2004	17-45
11	1993, 1995-2003	1990-2003	17-44
12	1993, 1995-2002	1990-2002	17-43
13	1993, 1995-2001	1990-2001	17-42
14	1993, 1995-2000	1990-2000	17-41
15	1993, 1995-1999	1990-1999	17-40
16	1993, 1995-1998	1990-1998	17-39

2.10.1 Web.com Tour ML Results

Table 2.5: Web.com Tour ML Earnings Results (Local Linear)

	All									Young		Old	
	year	τ	se	pval	β_l	β_r	N_l	N_r	h_l	h_r	τ	pval	τ	pval
	1	0.6454	0.1054	0.0000	12.90	12.27	319	879	-17.40	39.42	0.6161	0.0002	0.6293	0.0023
	2	0.2922	0.1583	0.0651	12.43	12.16	334	1066	-18.60	55.24	0.3996	0.1966	0.1420	0.5529
	3	0.2041	0.2074	0.3254	12.27	12.11	297	831	-17.84	47.64	0.3083	0.4059	0.0546	0.8495
	4	-0.0616	0.3832	0.8724	12.00	12.14	289	878	-20.71	60.67	-0.2089	0.6705	0.0305	0.9289
	5	0.1542	0.3509	0.6605	12.11	12.00	258	631	-22.82	48.25	-0.0511	0.9163	0.3147	0.2865
	6	-0.2416	0.3345	0.4704	11.83	12.07	228	657	-21.45	60.95	-0.4491	0.4269	0.0853	0.7563
	7	-0.5790	0.2657	0.0296	11.32	12.07	198	662	-21.32	75.47	-0.5690	0.1418	-0.6151	0.1818
World Earnings	8	-0.0310	0.2963	0.9167	11.79	12.06	167	505	-21.70	61.10	-0.4248	0.3547	0.3870	0.3952
	9	-0.1296	0.3483	0.7100	11.50	11.91	148	449	-22.19	68.57	-0.1752	0.6688	-0.5735	0.2287
	10	-0.4741	0.6626	0.4746	10.94	11.66	121	327	-22.51	56.45	-0.4586	0.5929	-1.4823	0.0109
	11	-0.3298	0.7944	0.6783	11.24	11.70	96	272	-23.19	57.92	-0.6596	0.4872	-0.6413	0.4118
	12	0.1159	0.4973	0.8159	11.38	11.10	72	187	-24.20	44.40	-0.2868	0.6022	0.4140	0.5926
	13	-1.4009	0.5143	0.0069	09.95	11.25	62	201	-26.06	64.67	-2.1585	0.0001	0.9948	0.3485
	14	-0.5172	1.2078	0.6689	10.40	10.70	41	169	-20.75	62.08	-1.9158	0.1148	4.4563	0.0330
	15	1.0357	1.2348	0.4029	11.45	10.36	32	132	-23.83	62.64	0.9474	0.4676	3.1077	0.1281
	16	-1.0194	1.2504	0.4164	09.86	10.90	28	113	-25.04	84.32	-2.0151	0.1084	-	-
	1	0.5454	0.1847	0.0032	12.81	12.28	363	391	-22.07	54.27	0.4687	0.0288	0.6067	0.0227
	2	0.2597	0.2092	0.2149	12.76	12.53	268	489	-23.11	52.83	0.5000	0.1562	-0.0263	0.9258
	3	0.1860	0.2404	0.4394	12.84	12.64	226	533	-21.71	71.34	0.2786	0.4193	0.0599	0.8000
	4	-0.0599	0.2546	0.8141	12.79	12.91	200	313	-21.38	40.18	-0.3184	0.3414	0.4232	0.1941
	5	-0.0592	0.5004	0.9059	12.73	12.87	181	363	-26.12	54.44	-0.3141	0.5721	0.3588	0.5560
	6	0.2326	0.2922	0.4263	12.97	12.74	156	393	-26.21	71.38	-0.0974	0.8150	1.0815	0.1828
	7	0.4128	0.3226	0.2012	13.15	12.78	141	390	-23.00	81.05	0.4667	0.4005	0.3684	0.6753
PGA TOUR Earnings	8	0.2924	0.4955	0.5553	12.79	12.64	118	358	-25.95	83.81	0.0415	0.9377	0.7884	0.3914
	9	-0.0556	0.5906	0.9250	12.43	12.65	97	202	-24.94	51.23	-0.1052	0.8827	-1.1891	0.2120
	10	0.1674	0.4833	0.7293	12.23	12.34	86	211	-24.89	66.99	-0.5197	0.2415	0.3484	0.7510
	11	-0.5759	0.5051	0.2553	11.79	12.38	65	208	-25.28	82.09	-0.9468	0.0791	-0.0674	0.9471
	12	0.0849	0.9286	0.9272	12.10	11.83	51	149	-25.99	69.61	0.2015	0.7804	0.0571	0.9859
	13	0.2624	0.8898	0.7685	12.15	11.87	47	97	-23.85	55.52	-0.2519	0.7676	2.1498	0.4302
	14	-0.1540	0.8461	0.8559	11.60	11.73	37	78	-19.93	59.92	-0.7361	0.3505	2.7346	0.2212
	15	-1.6453	0.9731	0.0943	10.46	12.28	26	73	-31.94	68.99	-0.8414	0.4947	-3.0135	0.0969
	16	-2.0794	1.0816	0.0589	10.92	12.92	21	52	-30.50	74.34	-1.7131	0.2922	-	-
	1	-2.0493	0.1911	0.0000	09.32	11.39	126	977	-15.21	48.81	-2.4917	0.0000	-1.4539	0.0015
	2	0.0891	0.2394	0.7099	11.03	10.97	184	1038	-20.66	73.32	-0.0418	0.9160	0.1762	0.4842
	3	-0.5069	0.1937	0.0091	10.28	10.82	161	602	-19.80	47.93	-0.8014	0.0471	-0.1984	0.5353
	4	-0.2944	0.3822	0.4414	10.31	10.71	137	784	-22.28	84.76	-0.3793	0.4095	-0.2737	0.5407
	5	0.2253	0.2850	0.4294	10.68	10.46	138	620	-24.94	80.12	-0.1744	0.7279	0.5968	0.0267
	6	-0.0785	0.2903	0.7871	10.45	10.55	110	477	-25.69	73.36	-0.8063	0.1688	0.5731	0.1146
	7	-0.8230	0.2241	0.0003	09.57	10.42	89	298	-21.61	48.16	-1.4545	0.0000	-0.0008	0.9986
Web.com Earnings	8	0.1267	0.2541	0.6183	10.39	10.32	74	369	-28.75	85.80	0.2793	0.5435	-0.3505	0.5549
	9	-0.1790	0.2796	0.5224	10.33	10.53	75	293	-24.06	83.76	-0.4315	0.3478	0.1833	0.6886
	10	-0.9349	0.4310	0.0309	09.43	10.44	59	245	-29.21	82.17	-0.3995	0.6039	-1.9046	0.0000
	11	0.1140	0.6621	0.8635	10.19	10.04	54	152	-26.97	54.79	-0.4354	0.5456	0.6452	0.4742
	12	-0.1292	0.7186	0.8576	09.61	09.72	41	119	-27.91	46.61	-0.9006	0.4040	1.5928	0.0199
	13	-1.2144	0.5184	0.0204	08.93	10.22	38	127	-34.69	71.41	-1.7462	0.0152	1.6979	0.0009
	14	-0.5429	1.0878	0.6186	09.02	09.44	18	104	-28.38	55.06	-0.9719	0.3639	-	-
	15	1.6423	0.7242	0.0274	10.36	08.37	14	46	-41.11	28.30	1.4673	0.3553	-	-
	16	1.6032	0.6426	0.0155	09.66	08.46	13	52	-35.92	46.60	-	-	-	-

Notes: *τ denotes the estimated average treatment effect. β denotes the estimated limits on each side of the treatment threshold. N denotes the number of observations. h denotes the length of the bandwidth. Young golfers are less than or equal to 30 years. See Section 2.5.1 for estimation details.*

Table 2.6: Web.com Tour ML Events Results (Local Linear)

	All									Young		Old	
	year	τ	se	pval	β_l	β_r	N_l	N_r	h_l	h_r	τ	pval	τ	pval
	1	16.6454	1.0481	0.0000	24.93	08.39	283	798	-13.72	35.27	14.8916	0.0000	18.3760	0.0000
	2	1.1913	1.1184	0.2870	11.90	10.92	283	1065	-14.09	47.04	2.0856	0.2471	-0.0444	0.9768
	3	3.7145	2.1895	0.0900	14.15	10.48	328	1079	-19.82	50.00	4.5916	0.1123	2.0221	0.4072
	4	0.9559	1.9577	0.6254	10.72	10.13	277	1030	-16.42	50.08	0.9535	0.6832	0.1937	0.9302
	5	0.6066	1.7959	0.7356	09.75	09.55	275	828	-17.70	42.46	0.6084	0.8545	-0.1547	0.8915
	6	-0.3807	1.4286	0.7899	08.33	08.83	257	878	-17.57	47.02	1.2550	0.6892	-2.5491	0.0396
	7	-1.1796	1.0442	0.2589	06.80	08.57	234	813	-16.85	46.20	-0.4845	0.8374	-2.0219	0.1992
PGA TOUR Events	8	-0.8666	0.9532	0.3634	06.62	08.21	217	896	-17.44	53.67	-1.9598	0.4291	0.3514	0.8302
	9	-0.7465	1.3386	0.5772	04.76	06.60	200	922	-16.69	59.29	-0.1896	0.9354	-1.7653	0.1782
	10	-0.4511	1.6693	0.7871	03.95	05.48	185	1004	-18.16	69.26	-2.4866	0.2609	-0.3589	0.8048
	11	-1.3546	1.3015	0.2982	02.84	04.88	169	764	-20.13	55.80	-0.8296	0.7312	-2.4761	0.0361
	12	-1.6450	1.1937	0.1686	02.04	04.07	149	658	-17.87	52.37	-1.8692	0.4313	-1.5461	0.2010
	13	-1.5234	1.1000	0.1665	01.61	03.60	134	675	-18.02	57.59	-2.4350	0.1559	-0.7223	0.5821
	14	0.3699	1.4878	0.8037	02.68	02.57	119	679	-18.31	64.22	-0.5214	0.8308	1.1280	0.3656
	15	-0.0026	1.4362	0.9985	02.08	02.04	104	708	-18.81	73.56	-0.8401	0.6924	0.7894	0.5549
	16	-1.2987	1.4918	0.3843	00.98	02.04	89	578	-18.09	66.77	-1.8805	0.4397	-1.1555	0.4181
	1	-16.2806	1.0495	0.0000	01.47	17.75	329	548	-18.02	23.83	-15.3831	0.0000	-16.8311	0.0000
	2	-1.0818	1.1839	0.3610	11.56	12.56	299	1634	-15.48	73.35	-3.7462	0.0080	1.3517	0.3639
	3	-3.5867	0.9972	0.0003	07.31	10.88	320	1418	-18.54	65.71	-6.4468	0.0000	-0.2472	0.8796
	4	-0.9861	1.0816	0.3620	08.06	08.97	277	1671	-15.87	81.60	-2.4429	0.0453	1.1665	0.5045
	5	-0.3848	1.0546	0.7152	06.89	07.35	275	1532	-18.30	79.37	-2.0061	0.1658	1.1413	0.5217
	6	0.3552	0.9728	0.7150	07.04	06.63	251	1509	-16.75	82.31	-1.5894	0.3692	2.7566	0.1784
	7	-1.1646	0.5674	0.0403	05.03	06.13	229	1340	-15.85	77.01	-3.0193	0.0105	1.0155	0.3700
Web.com Events	8	-1.1148	1.0673	0.2965	04.55	05.61	221	980	-17.68	59.16	-1.3900	0.2352	-0.6597	0.7036
	9	0.7581	1.3321	0.5694	06.10	05.27	209	859	-21.84	54.74	0.8493	0.7935	0.4544	0.7473
	10	-1.6591	1.2813	0.1956	03.48	05.21	189	918	-20.44	63.25	-2.8211	0.1382	-1.2982	0.4414
	11	-2.1631	1.0845	0.0463	01.70	04.06	167	1115	-18.14	83.35	-2.3907	0.1090	-2.1479	0.1849
	12	-1.5040	0.9373	0.1089	01.78	03.57	149	782	-19.12	62.15	-1.1613	0.4393	-1.8206	0.2607
	13	-1.6157	0.9708	0.0964	01.84	03.81	134	752	-19.81	65.46	-1.3596	0.5047	-1.8685	0.2122
	14	-1.9173	0.7624	0.0121	01.27	03.35	119	818	-18.92	77.95	-1.0829	0.4029	-2.0318	0.0483
	15	-0.2886	0.7722	0.7088	00.75	01.26	104	394	-21.75	39.74	-0.3625	0.8229	-0.8188	0.0808
	16	1.2232	1.0950	0.2648	01.51	00.54	89	232	-22.03	25.56	0.8500	0.5042	1.3585	0.5529
	1	0.5399	0.9677	0.5770	26.44	25.99	283	821	-14.33	36.37	-0.9287	0.3952	1.8969	0.0188
	2	0.2901	1.0703	0.7864	23.43	23.22	299	1129	-15.31	49.74	-0.8896	0.5082	1.4374	0.3253
	3	0.1016	2.2243	0.9636	21.35	21.30	311	1058	-18.08	48.92	-1.7859	0.4906	1.8233	0.4358
	4	0.4421	1.9857	0.8238	18.85	18.71	285	1071	-16.96	51.98	-0.4624	0.7902	1.3956	0.5285
	5	0.4430	2.1664	0.8380	16.63	16.69	275	1001	-18.45	50.78	-0.5775	0.8259	1.1229	0.5643
	6	0.1318	1.6140	0.9350	15.39	15.37	257	931	-18.33	49.75	0.0125	0.9955	0.6669	0.7239
	7	-2.1806	1.0804	0.0438	11.69	14.45	229	898	-15.94	50.77	-2.3132	0.2888	-0.9579	0.6580
US Events	8	-2.1687	1.3979	0.1211	10.85	13.81	217	843	-16.70	51.25	-2.9869	0.2130	-0.2316	0.9346
	9	0.0515	1.4295	0.9713	10.75	11.78	202	836	-17.91	54.14	0.8954	0.7395	-1.1892	0.5172
	10	-2.1237	1.9895	0.2860	07.48	10.62	187	837	-19.52	58.17	-4.8778	0.0696	-1.3336	0.5464
	11	-3.7440	2.0317	0.0657	04.52	08.92	163	841	-18.74	62.65	-3.1277	0.2306	-5.2477	0.0066
	12	-3.5389	1.7352	0.0417	03.88	07.80	145	780	-18.79	63.59	-3.2144	0.2361	-3.9685	0.0199
	13	-3.1710	1.5801	0.0451	03.70	07.38	129	597	-19.90	53.07	-3.4468	0.0647	-3.3396	0.0724
	14	-1.6304	1.9301	0.3985	04.14	05.99	115	736	-18.42	72.98	-2.0138	0.4661	-1.2390	0.5196
	15	-0.3252	1.8816	0.8629	02.98	03.25	100	376	-20.80	40.05	-0.7585	0.7703	-0.0917	0.9561
	16	0.1391	2.6092	0.9575	02.84	02.44	84	327	-20.89	39.02	-0.7254	0.7809	0.6279	0.8652

Notes: *See notes to Table 2.5.*

Table 2.7: Web.com Tour ML Employment Results (Local Linear)

	All							Young		Old	
	year	τ	se	pval	$\beta_l \mid \beta_r$	$N_l \mid N_r$	$h_l \mid h_r$	τ	pval	τ	pval
	1	0.0017	0.0046	0.7018	$1.000 \mid 0.999$	$198 \mid 366$	$-08.67 \mid 16.11$	-0.0005	0.9352	0.0093	0.2942
	2	0.0063	0.0257	0.8053	$0.980 \mid 0.974$	$283 \mid 1065$	$-14.29 \mid 46.64$	-0.0158	0.6593	0.0231	0.6054
	3	0.0000	0.0551	0.9997	$0.908 \mid 0.908$	$311 \mid 1039$	$-17.76 \mid 48.01$	-0.0186	0.7007	0.0242	0.7337
	4	0.0980	0.0490	0.0458	$0.915 \mid 0.825$	$269 \mid 909$	$-15.28 \mid 44.41$	0.0467	0.2630	0.1214	0.1007
	5	0.0013	0.0588	0.9829	$0.775 \mid 0.784$	$275 \mid 944$	$-17.50 \mid 47.76$	-0.0075	0.9371	0.0108	0.8854
	6	0.0141	0.0521	0.7864	$0.757 \mid 0.741$	$257 \mid 878$	$-17.96 \mid 47.33$	0.0878	0.1507	-0.0615	0.4985
	7	0.0118	0.0477	0.8047	$0.665 \mid 0.664$	$229 \mid 759$	$-15.94 \mid 43.39$	0.0144	0.8226	0.0398	0.7035
+ World Earnings	8	-0.0746	0.0597	0.2117	$0.545 \mid 0.636$	$217 \mid 745$	$-16.73 \mid 45.46$	-0.0585	0.5363	-0.0462	0.6648
	9	0.0158	0.0846	0.8517	$0.554 \mid 0.571$	$202 \mid 746$	$-17.60 \mid 47.84$	0.0288	0.8614	-0.0138	0.8868
	10	-0.0828	0.0668	0.2154	$0.428 \mid 0.547$	$183 \mid 738$	$-17.82 \mid 50.68$	-0.2177	0.0054	-0.0091	0.9375
	11	-0.1856	0.0759	0.0147	$0.258 \mid 0.458$	$161 \mid 803$	$-17.29 \mid 59.60$	-0.2171	0.0284	-0.1942	0.0143
	12	-0.2201	0.0616	0.0004	$0.190 \mid 0.423$	$145 \mid 602$	$-17.49 \mid 49.24$	-0.2439	0.0504	-0.2187	0.0076
	13	0.0010	0.0658	0.9883	$0.373 \mid 0.391$	$129 \mid 620$	$-19.95 \mid 55.12$	0.1539	0.0463	-0.1646	0.0942
	14	-0.0762	0.0739	0.3031	$0.268 \mid 0.353$	$115 \mid 574$	$-18.98 \mid 56.41$	0.0358	0.7512	-0.2389	0.0062
	15	-0.1071	0.0967	0.2684	$0.153 \mid 0.261$	$100 \mid 518$	$-19.70 \mid 55.62$	-0.0818	0.5047	-0.0671	0.5740
	16	-0.0580	0.1155	0.6156	$0.120 \mid 0.162$	$84 \mid 391$	$-21.88 \mid 47.29$	-0.0412	0.7393	-0.0708	0.5875
	1	0.5762	0.0318	0.0000	$0.999 \mid 0.422$	$235 \mid 931$	$-11.12 \mid 41.16$	0.5390	0.0000	0.5920	0.0000
	2	0.0871	0.0400	0.0298	$0.606 \mid 0.522$	$329 \mid 997$	$-17.54 \mid 43.58$	0.0819	0.2093	0.0933	0.2706
	3	0.0966	0.0701	0.1686	$0.609 \mid 0.513$	$328 \mid 911$	$-19.92 \mid 41.85$	0.1424	0.1420	0.0440	0.6203
	4	0.0841	0.1169	0.4718	$0.533 \mid 0.457$	$293 \mid 1051$	$-17.84 \mid 50.90$	0.0846	0.5915	0.0379	0.7133
	5	0.0480	0.0798	0.5478	$0.476 \mid 0.441$	$268 \mid 905$	$-17.47 \mid 46.33$	0.0945	0.4548	-0.0158	0.8160
	6	-0.0336	0.0608	0.5804	$0.387 \mid 0.421$	$251 \mid 803$	$-17.22 \mid 42.96$	0.0484	0.6530	-0.1727	0.0029
	7	-0.0767	0.0446	0.0862	$0.316 \mid 0.412$	$229 \mid 777$	$-16.24 \mid 44.30$	-0.0420	0.5396	-0.1139	0.0763
+ PGA TOUR Earnings	8	-0.1085	0.0504	0.0317	$0.269 \mid 0.401$	$217 \mid 826$	$-16.89 \mid 49.51$	-0.1626	0.0087	-0.0547	0.5437
	9	-0.0836	0.0639	0.1908	$0.201 \mid 0.325$	$199 \mid 806$	$-16.99 \mid 51.60$	-0.0398	0.6924	-0.1271	0.0221
	10	-0.0391	0.0631	0.5358	$0.196 \mid 0.275$	$181 \mid 963$	$-17.23 \mid 66.98$	-0.1001	0.3768	-0.0468	0.3000
	11	-0.0708	0.0783	0.3661	$0.161 \mid 0.251$	$164 \mid 894$	$-19.59 \mid 67.02$	-0.0466	0.6917	-0.1376	0.0168
	12	-0.1253	0.0678	0.0651	$0.101 \mid 0.236$	$145 \mid 687$	$-18.35 \mid 55.70$	-0.1006	0.3333	-0.1494	0.0604
	13	-0.0727	0.0791	0.3581	$0.132 \mid 0.222$	$129 \mid 608$	$-18.50 \mid 54.35$	-0.0717	0.5076	-0.0749	0.3615
	14	0.0190	0.1038	0.8547	$0.183 \mid 0.166$	$115 \mid 497$	$-19.44 \mid 48.23$	0.0311	0.8408	-0.0173	0.8269
	15	0.0122	0.1059	0.9086	$0.161 \mid 0.137$	$100 \mid 688$	$-19.60 \mid 75.74$	-0.0355	0.7267	0.0890	0.4844
	16	-0.0545	0.0878	0.5350	$0.060 \mid 0.102$	$84 \mid 603$	$-19.50 \mid 74.53$	-0.1431	0.1557	0.0070	0.9597
	1	-0.4351	0.0529	0.0000	$0.433 \mid 0.868$	$319 \mid 502$	$-17.24 \mid 21.68$	-0.2627	0.0003	-0.5776	0.0000
	2	0.0184	0.0507	0.7172	$0.686 \mid 0.661$	$299 \mid 1218$	$-14.92 \mid 53.63$	-0.0918	0.1566	0.1091	0.1420
	3	-0.1661	0.0474	0.0005	$0.482 \mid 0.645$	$284 \mid 1439$	$-14.87 \mid 66.94$	-0.2381	0.0025	-0.0592	0.3726
	4	0.0478	0.0447	0.2856	$0.588 \mid 0.531$	$255 \mid 1569$	$-14.47 \mid 77.07$	-0.0797	0.3049	0.1785	0.0175
	5	0.0198	0.0512	0.6988	$0.467 \mid 0.446$	$282 \mid 1078$	$-18.77 \mid 54.99$	0.0303	0.7689	0.0059	0.9298
	6	0.0552	0.0554	0.3193	$0.459 \mid 0.398$	$257 \mid 1331$	$-18.22 \mid 71.60$	0.0376	0.7330	0.0590	0.5286
	7	0.0264	0.0369	0.4742	$0.432 \mid 0.397$	$244 \mid 1289$	$-18.51 \mid 74.38$	-0.0069	0.9413	0.0936	0.1770
+ Web.com Earnings	8	-0.0235	0.0523	0.6542	$0.335 \mid 0.354$	$221 \mid 859$	$-17.51 \mid 52.09$	-0.0657	0.5086	0.0423	0.4838
	9	0.0579	0.0931	0.5339	$0.377 \mid 0.321$	$205 \mid 836$	$-19.32 \mid 53.80$	0.0874	0.6220	0.0386	0.5906
	10	-0.0282	0.0712	0.6925	$0.289 \mid 0.323$	$187 \mid 879$	$-19.89 \mid 61.42$	-0.1121	0.3168	0.0395	0.6788
	11	-0.1195	0.0608	0.0496	$0.151 \mid 0.276$	$161 \mid 934$	$-17.18 \mid 70.09$	-0.0928	0.4691	-0.1425	0.0683
	12	-0.1015	0.0949	0.2854	$0.130 \mid 0.239$	$145 \mid 639$	$-18.57 \mid 52.45$	-0.0278	0.8106	-0.1689	0.0671
	13	-0.0207	0.0754	0.7840	$0.192 \mid 0.222$	$129 \mid 805$	$-20.54 \mid 71.65$	0.1797	0.2412	-0.2079	0.0241
	14	-0.0365	0.0548	0.5059	$0.210 \mid 0.254$	$115 \mid 603$	$-19.43 \mid 59.45$	0.0499	0.7133	-0.1695	0.0616
	15	-0.1220	0.0674	0.0706	$0.048 \mid 0.181$	$100 \mid 642$	$-19.07 \mid 69.91$	-0.0719	0.4299	-0.1113	0.2686
	16	0.0400	0.0887	0.6524	$0.098 \mid 0.064$	$84 \mid 256$	$-21.02 \mid 30.47$	0.0889	0.3466	-0.0248	0.8545

Notes: *"+ World Earnings" denotes positive world earnings. See notes to Table 2.5.*

Table 2.8: Web.com Tour ML Performance Results (Local Linear)

	All							Young		Old	
	year	τ	se	pval	$\beta_l \mid \beta_r$	$N_l \mid N_r$	$h_l \mid h_r$	τ	pval	τ	pval
Scoring Average	1	-0.0857	0.0650	0.1876	-0.036 | 0.037	299 | 1008	-15.07 | 44.68	-0.0755	0.5245	-0.0986	0.4636
	2	-0.1188	0.1075	0.2689	0.067 | 0.176	326 | 1300	-17.62 | 64.72	-0.2053	0.3487	0.0023	0.9897
	3	-0.0749	0.1079	0.4877	0.208 | 0.260	294 | 1200	-17.33 | 67.76	-0.1318	0.5947	0.0506	0.7596
	4	-0.1339	0.1645	0.4157	0.264 | 0.360	291 | 794	-20.01 | 48.46	-0.2137	0.4127	-0.0198	0.9365
	5	0.0379	0.1401	0.7869	0.350 | 0.313	244 | 706	-17.88 | 48.48	0.1689	0.6595	-0.1322	0.6782
	6	0.1896	0.1573	0.2284	0.438 | 0.222	220 | 683	-18.48 | 54.53	0.0974	0.7030	0.3141	0.4289
	7	0.1118	0.2178	0.6076	0.641 | 0.442	209 | 849	-20.76 | 84.79	0.0661	0.7629	0.2115	0.6702
	8	-0.2181	0.1893	0.2497	0.202 | 0.325	173 | 603	-19.92 | 64.46	-0.2063	0.4374	-0.2038	0.5607
	9	-0.1821	0.2317	0.4321	0.604 | 0.628	157 | 592	-20.54 | 79.70	-0.1847	0.5108	0.8261	0.0714
	10	0.1819	0.4099	0.6573	0.886 | 0.557	125 | 359	-20.70 | 49.95	0.2975	0.6554	0.6906	0.1310
	11	0.5579	0.4945	0.2599	1.122 | 0.471	100 | 357	-21.35 | 62.17	1.3939	0.1078	-0.6790	0.2524
	12	1.2637	0.4239	0.0031	2.347 | 1.032	92 | 265	-20.27 | 55.18	1.0850	0.0024	1.5192	0.1302
	13	0.9114	0.6471	0.1600	2.095 | 1.262	74 | 253	-24.08 | 64.40	0.3339	0.4512	3.3700	0.2381
	14	1.1808	0.5156	0.0227	2.007 | 1.035	53 | 240	-22.04 | 75.50	1.3787	0.0059	-1.0688	0.1444
	15	0.0563	0.6786	0.9340	1.803 | 1.638	44 | 135	-21.17 | 47.47	0.1095	0.8717	-1.4797	0.3910
	16	0.8740	0.4472	0.0522	2.230 | 1.014	35 | 154	-26.32 | 84.13	1.3324	0.0789	-1.3435	0.0480
Unadj SA	1	0.6659	0.0658	0.0000	0.298 | -0.379	299 | 1262	-15.30 | 57.42	0.6377	0.0000	0.7016	0.0000
	2	-0.0743	0.1233	0.5469	-0.122 | -0.055	345 | 1268	-19.79 | 63.46	-0.0556	0.7689	-0.0474	0.7130
	3	0.0701	0.0890	0.4308	0.127 | 0.044	303 | 1183	-18.14 | 66.69	0.0976	0.6063	0.1237	0.4683
	4	-0.1077	0.1388	0.4379	0.084 | 0.160	291 | 895	-20.20 | 55.43	-0.1952	0.2613	-0.0351	0.8855
	5	-0.0193	0.1094	0.8597	0.171 | 0.185	256 | 894	-19.94 | 62.94	0.0214	0.9257	-0.0988	0.7360
	6	0.0708	0.1391	0.6108	0.215 | 0.132	231 | 872	-19.97 | 75.30	0.0103	0.9480	0.1047	0.7276
	7	-0.0504	0.1759	0.7746	0.389 | 0.393	209 | 650	-21.35 | 60.11	-0.0188	0.9129	0.0029	0.9940
	8	-0.2409	0.2274	0.2899	0.028 | 0.243	175 | 615	-21.26 | 66.21	-0.2260	0.3857	-0.2329	0.4803
	9	-0.3030	0.2179	0.1648	0.303 | 0.505	157 | 518	-21.29 | 64.10	-0.2744	0.3101	0.4872	0.2739
	10	0.1948	0.3081	0.5274	0.689 | 0.381	125 | 439	-21.94 | 62.90	0.3029	0.5241	0.3968	0.3948
	11	0.5178	0.4094	0.2067	0.991 | 0.371	100 | 335	-22.27 | 59.07	1.4124	0.0385	-0.9599	0.1605
	12	1.0091	0.3655	0.0061	2.020 | 0.953	92 | 236	-21.34 | 47.35	1.0385	0.0293	0.9819	0.3006
	13	0.7232	0.6653	0.2778	1.762 | 1.060	74 | 274	-24.83 | 70.58	0.1158	0.8189	3.1077	0.2571
	14	1.0542	0.4427	0.0179	1.657 | 0.763	53 | 245	-22.50 | 76.53	1.0693	0.0209	-0.5866	0.4368
	15	0.1171	0.5480	0.8311	1.689 | 1.431	44 | 130	-21.48 | 43.89	0.0234	0.9590	-0.8526	0.5539
	16	0.5650	0.5739	0.3266	1.925 | 0.998	35 | 111	-27.24 | 55.18	0.9571	0.2905	-1.4960	0.2037

Notes: *See notes to Table 2.5.*

Table 2.9: Web.com Tour ML Total Earnings Effects (Local Linear)

	All						Young		Old	
	year	τ	se	pval	$\beta_l \mid \beta_r$	$N_l \mid N_r$	$h_l \mid h_r$	τ	pval	τ	pval
World Earnings	1	164,108	46,984	0.0005	508,470 | 352,465	299 | 798	-15.08 | 35.16	178,850	0.1149	138,916	0.0661
	2	123,350	66,545	0.0640	557,544 | 444,697	309 | 888	-16.39 | 39.29	150,944	0.3485	95,865	0.1579
	3	96,454	96,863	0.3195	502,938 | 414,562	293 | 1273	-16.19 | 59.48	82,117	0.6876	63,211	0.2887
	4	63,700	108,870	0.5586	511,699 | 469,796	285 | 1091	-17.23 | 53.03	16,716	0.9385	72,135	0.2021
	5	12,921	183,070	0.9437	409,626 | 420,021	275 | 1176	-17.75 | 59.52	-24,914	0.9487	7,108	0.9436
	6	-13,352	138,091	0.9230	458,829 | 465,133	263 | 912	-18.99 | 48.80	-68,999	0.8146	-23,303	0.7100
	7	249,269	136,343	0.0677	575,379 | 355,953	234 | 1240	-17.22 | 70.98	381,266	0.3123	35,559	0.6546
	8	-99,893	79,530	0.2093	235,180 | 363,997	213 | 1022	-16.26 | 61.53	-291,482	0.1018	42,973	0.5641
	9	170,497	123,876	0.1690	389,786 | 264,120	205 | 985	-18.76 | 63.65	360,737	0.2235	-77,260	0.3333
	10	-65,503	146,207	0.6542	127,360 | 245,221	177 | 1110	-15.39 | 77.66	-50,567	0.8466	-91,381	0.3346
	11	-28,091	113,257	0.8042	156,039 | 212,341	160 | 908	-16.45 | 67.93	-30,562	0.8945	-49,123	0.2559
	12	4,825	104,242	0.9631	130,453 | 134,585	145 | 908	-17.38 | 75.49	16,883	0.9402	-19,744	0.7474
	13	-100,711	79,403	0.2050	053,963 | 170,214	129 | 663	-18.06 | 58.83	-218,027	0.1806	-7,289	0.8442
	14	-86,447	83,336	0.2999	034,316 | 125,577	115 | 692	-16.54 | 67.77	-232,651	0.0950	34,890	0.6366
	15	-143,927	66,954	0.0319	-001,302 | 133,734	100 | 723	-16.39 | 80.20	-255,095	0.0082	-21,334	0.6126
	16	-210,646	115,867	0.0694	-069,067 | 120,011	84 | 728	-14.59 | 92.08	-296,932	0.0379	-106,962	0.2197
PGA TOUR Earnings	1	290,676	48,776	0.0000	489,968 | 206,563	299 | 709	-15.14 | 30.60	325,189	0.0035	236,024	0.0008
	2	111,846	59,387	0.0599	448,009 | 348,272	309 | 910	-15.73 | 40.15	186,521	0.2269	34,804	0.6081
	3	117,594	98,075	0.2307	452,767 | 342,940	293 | 1336	-16.17 | 62.24	127,777	0.5276	59,600	0.2724
	4	70,001	110,709	0.5273	460,795 | 412,685	285 | 1051	-16.98 | 50.80	37,709	0.8668	62,021	0.3463
	5	26,040	179,658	0.8848	364,702 | 361,698	275 | 1157	-17.64 | 59.17	12,274	0.9747	-645	0.9949
	6	5,430	137,571	0.9685	429,092 | 416,624	263 | 912	-19.40 | 48.87	-32,330	0.9141	-29,210	0.6537
	7	266,322	134,571	0.0480	549,344 | 313,587	234 | 1224	-17.05 | 70.48	409,870	0.2811	35,092	0.6526
	8	-87,074	78,887	0.2699	208,064 | 324,738	213 | 1084	-16.32 | 65.60	-271,877	0.1244	52,125	0.4749
	9	173,001	123,718	0.1622	365,003 | 234,955	205 | 1160	-18.83 | 75.82	365,406	0.2160	-69,503	0.3904
	10	-41,520	146,358	0.7767	117,271 | 211,022	177 | 1069	-15.31 | 74.79	-29,334	0.9105	-67,976	0.4754
	11	-17,983	110,326	0.8706	144,910 | 190,190	160 | 983	-16.32 | 73.76	-12,234	0.9574	-46,308	0.3167
	12	12,995	102,984	0.8996	123,909 | 117,681	145 | 1004	-17.12 | 83.15	51,860	0.8186	-23,232	0.7182
	13	-90,149	78,135	0.2489	048,963 | 153,816	129 | 716	-17.92 | 64.14	-198,276	0.2274	-14,356	0.7188
	14	-79,928	82,864	0.3351	031,054 | 116,023	115 | 681	-16.60 | 67.12	-217,872	0.1171	35,492	0.6347
	15	-134,553	65,822	0.0412	-000,430 | 124,570	100 | 751	-16.33 | 82.91	-239,804	0.0146	-23,112	0.5883
	16	-213,797	115,861	0.0654	-075,726 | 115,959	84 | 728	-14.58 | 92.06	-291,123	0.0441	-124,119	0.1762
Web.com Earnings	1	-133,080	10,478	0.0000	010,382 | 144,074	309 | 1264	-16.05 | 55.82	-162,677	0.0000	-111,958	0.0000
	2	8,614	20,728	0.6778	091,082 | 083,328	319 | 1942	-16.57 | 87.13	-19,960	0.4090	39,579	0.0384
	3	-21,561	6,526	0.0010	040,616 | 062,938	293 | 1730	-15.61 | 81.21	-35,328	0.0236	-5,906	0.6064
	4	-436	9,250	0.9624	050,817 | 051,914	269 | 1414	-14.88 | 69.25	-12,055	0.3370	17,796	0.1884
	5	307	13,066	0.9813	038,933 | 040,192	282 | 1609	-18.59 | 83.39	-5,063	0.8067	10,700	0.2719
	6	-1,288	5,987	0.8297	035,253 | 036,199	245 | 1598	-15.82 | 87.18	-11,358	0.3717	10,340	0.3150
	7	-6,758	5,441	0.2144	023,681 | 031,209	229 | 1190	-16.30 | 68.39	-19,511	0.0769	12,639	0.0567
	8	-3,590	6,021	0.5510	024,174 | 028,150	225 | 1181	-18.87 | 71.51	-1,238	0.9044	-6,976	0.4880
	9	-1,151	6,764	0.8648	024,480 | 026,534	205 | 1118	-18.86 | 72.99	153	0.9908	-1,529	0.8900
	10	-14,879	6,248	0.0174	007,906 | 023,738	183 | 1193	-18.02 | 83.97	-5,018	0.6715	-26,375	0.0086
	11	-7,935	6,925	0.2521	009,017 | 017,607	164 | 1044	-19.94 | 79.34	-15,403	0.0356	-3,167	0.7995
	12	-6,806	5,879	0.2473	005,092 | 012,927	145 | 780	-19.39 | 63.73	-12,617	0.2164	-1,685	0.8745
	13	-7,081	5,248	0.1776	004,784 | 012,371	129 | 841	-20.83 | 76.00	-13,946	0.1115	-275	0.9753
	14	-556	4,191	0.8944	005,553 | 006,302	115 | 514	-19.64 | 50.45	1,653	0.8029	-2,199	0.5000
	15	-1,370	3,714	0.7123	000,967 | 002,993	100 | 473	-21.20 | 50.87	-5,575	0.4182	1,410	0.4841
	16	6,747	6,326	0.2870	006,560 | -000,168	84 | 222	-22.86 | 26.43	944	0.8345	18,814	0.3578

Notes: *These regressions are estimated in earnings levels rather than logs. See notes to Table 2.5.*

2.10.2 Q School Results

Table 2.10: Q School Earnings Results (Local Randomization)

	All						Young		Old	
	year	τ	pval	β_l | β_r	N_l | N_r	h_l | h_r	τ	pval	τ	pval
World Earnings	1	0.9366	0.0000	12.56 | 11.62	120 | 108	-0.5 | 0.5	0.7852	0.0000	1.0999	0.0000
	2	0.1554	0.5120	12.21 | 12.05	98 | 92	-0.5 | 0.5	-0.0387	0.9110	0.3467	0.2760
	3	0.0619	0.8180	12.05 | 11.99	88 | 86	-0.5 | 0.5	0.3337	0.3370	-0.1836	0.6110
	4	0.0544	0.8500	12.07 | 12.02	79 | 76	-0.5 | 0.5	0.0585	0.8750	0.1502	0.6800
	5	0.0912	0.7860	12.03 | 11.94	65 | 75	-0.5 | 0.5	0.1710	0.6870	0.1847	0.7380
	6	-0.0037	0.9920	11.98 | 11.98	51 | 74	-0.5 | 0.5	-0.2002	0.6930	0.3728	0.4530
	7	-0.2949	0.4760	11.68 | 11.98	49 | 64	-0.5 | 0.5	-0.5200	0.3270	0.1877	0.7320
	8	-0.4569	0.3090	11.73 | 12.19	36 | 55	-0.5 | 0.5	-0.5563	0.2870	0.0462	0.9510
	9	-0.8774	0.0560	11.65 | 12.53	35 | 45	-0.5 | 0.5	-0.9088	0.1280	-0.6902	0.2600
	10	-0.0265	0.9440	11.89 | 11.92	34 | 36	-0.5 | 0.5	-0.3291	0.5550	0.5314	0.3000
	11	-0.4508	0.3240	11.61 | 12.06	31 | 27	-0.5 | 0.5	-0.2771	0.6230	-1.0473	0.1420
	12	0.3557	0.4910	12.01 | 11.66	29 | 20	-0.5 | 0.5	0.9986	0.1340	-1.1652	0.1000
	13	0.4639	0.4860	11.26 | 10.80	25 | 21	-0.5 | 0.5	1.3113	0.1050	-1.1607	0.2530
	14	-0.3103	0.6350	11.30 | 11.61	18 | 19	-0.5 | 0.5	0.0944	0.9060	-1.4516	0.1420
	15	0.2965	0.6790	11.38 | 11.08	18 | 13	-0.5 | 0.5	0.2574	0.7690	-0.3528	1.0000
	16	-0.1799	0.8490	10.59 | 10.77	12 | 13	-0.5 | 0.5	-0.1871	0.8430	-1.9041	0.1200
PGA TOUR Earnings	1	1.2708	0.0000	12.50 | 11.22	32 | 107	-0.5 | 0.5	2.0551	0.0000	1.0126	0.0020
	2	0.3841	0.1670	12.81 | 12.43	50 | 57	-0.5 | 0.5	0.3386	0.3610	0.5202	0.1520
	3	0.1528	0.6320	12.54 | 12.39	50 | 52	-0.5 | 0.5	-0.1447	0.7390	0.3168	0.4640
	4	0.0613	0.8650	12.31 | 12.25	42 | 50	-0.5 | 0.5	0.1241	0.7950	0.1632	0.7470
	5	-0.5349	0.2000	12.06 | 12.60	39 | 48	-0.5 | 0.5	0.2246	0.6740	-1.1611	0.0710
	6	0.0257	0.9440	12.73 | 12.71	33 | 43	-0.5 | 0.5	0.0171	0.9580	0.3286	0.6250
	7	-0.1944	0.6340	12.73 | 12.92	29 | 33	-0.5 | 0.5	0.0504	0.9150	0.0968	0.9090
	8	0.3249	0.4810	12.88 | 12.56	24 | 27	-0.5 | 0.5	0.4464	0.4450	0.4310	0.6600
	9	0.4016	0.3870	13.20 | 12.80	23 | 22	-0.5 | 0.5	0.4249	0.4560	0.5675	0.3850
	10	0.1822	0.7290	12.49 | 12.31	22 | 19	-0.5 | 0.5	0.0700	0.9250	0.7518	0.1920
	11	-0.2625	0.6340	12.18 | 12.44	19 | 16	-0.5 | 0.5	-0.4112	0.5100	0.1868	0.8120
	12	0.1443	0.8480	12.34 | 12.20	16 | 13	-0.5 | 0.5	0.4900	0.5360	-1.0818	0.4950
	13	0.1683	0.8460	12.12 | 11.95	12 | 13	-0.5 | 0.5	0.6156	0.5050	-1.0835	0.7950
	14	0.2906	0.7480	12.24 | 11.94	9 | 11	-0.5 | 0.5	0.7466	0.4420	-2.4553	0.6740
	15	-0.1309	0.9120	12.31 | 12.44	8 | 6	-0.5 | 0.5	-0.3851	0.7140	-	-
	16	1.4949	0.2920	12.85 | 11.35	7 | 4	-0.5 | 0.5	1.2764	0.4200	-	-
Web.com Earnings	1	-1.9313	0.0000	09.29 | 11.22	110 | 54	-0.5 | 0.5	-2.0923	0.0000	-1.8357	0.0000
	2	-0.5041	0.0710	10.16 | 10.66	68 | 56	-0.5 | 0.5	-0.2989	0.4800	-0.7006	0.0860
	3	-0.2093	0.5040	10.46 | 10.67	50 | 50	-0.5 | 0.5	0.3848	0.3080	-0.7286	0.1150
	4	-0.1942	0.5600	10.42 | 10.62	49 | 45	-0.5 | 0.5	-0.1664	0.7310	-0.1104	0.8010
	5	0.1873	0.6130	10.37 | 10.18	36 | 42	-0.5 | 0.5	-0.1007	0.8690	0.4802	0.3480
	6	-0.1046	0.8000	09.90 | 10.01	27 | 41	-0.5 | 0.5	-0.6554	0.2200	0.5888	0.2920
	7	0.2327	0.5680	09.89 | 09.66	25 | 33	-0.5 | 0.5	0.4364	0.4130	-0.0399	0.9490
	8	0.0285	0.9570	10.35 | 10.32	16 | 28	-0.5 | 0.5	-0.0412	0.9530	-0.0050	0.9970
	9	-0.3773	0.4960	10.04 | 10.42	15 | 21	-0.5 | 0.5	-0.4020	0.6150	-0.3626	0.6010
	10	0.1495	0.7770	10.38 | 10.23	15 | 20	-0.5 | 0.5	-0.1665	0.8270	0.3169	0.6790
	11	-0.4823	0.3900	09.64 | 10.12	14 | 16	-0.5 | 0.5	-0.2064	0.7830	-0.7572	0.3820
	12	0.0003	1.0000	10.03 | 10.03	17 | 12	-0.5 | 0.5	0.4235	0.5670	-1.3029	0.1220
	13	0.3653	0.5840	09.61 | 09.24	16 | 11	-0.5 | 0.5	0.7821	0.2970	-0.1047	0.9020
	14	-0.2319	0.6440	10.37 | 10.60	11 | 10	-0.5 | 0.5	0.2219	0.7820	-0.9274	0.3260
	15	0.7631	0.2910	10.59 | 09.83	10 | 8	-0.5 | 0.5	0.9371	0.3010	0.2154	1.0000
	16	0.3189	0.7290	09.94 | 09.62	6 | 7	-0.5 | 0.5	1.0347	0.2990	-1.5835	0.6670

Notes: *The local randomization software does not compute a standard error for the ATE. See notes to Table 2.5.*

Table 2.11: Q School Events Results (Local Randomization)

	All						Young		Old	
	year	τ	pval	$\beta_l \mid beta_r$	$N_l \mid N_r$	$h_l \mid h_r$	τ	pval	τ	pval
	1	21.0596	0.0000	23.10 | 02.04	121 | 109	-0.5 | 0.5	22.2344	0.0000	19.7173	0.0000
	2	2.9227	0.0550	11.14 | 08.21	121 | 109	-0.5 | 0.5	1.0290	0.6460	4.9229	0.0230
	3	1.0161	0.5450	09.89 | 08.87	116 | 106	-0.5 | 0.5	0.6667	0.7760	1.3272	0.5580
	4	1.7500	0.2750	09.00 | 07.25	108 | 103	-0.5 | 0.5	0.9401	0.6940	2.7142	0.1770
	5	-0.3515	0.8410	07.86 | 08.21	104 | 100	-0.5 | 0.5	-1.7153	0.5030	1.2838	0.5100
	6	1.7012	0.3100	08.98 | 07.28	97 | 98	-0.5 | 0.5	0.5157	0.8650	3.1923	0.1260
	7	0.7189	0.6840	07.61 | 06.89	95 | 88	-0.5 | 0.5	-0.0521	0.9830	1.8889	0.3210
PGA TOUR Events	8	1.2137	0.4580	06.85 | 05.63	90 | 85	-0.5 | 0.5	0.3918	0.8860	2.4434	0.1930
	9	0.4383	0.7720	05.92 | 05.48	87 | 76	-0.5 | 0.5	0.3514	0.8960	1.1301	0.4390
	10	0.1817	0.9130	05.17 | 04.99	79 | 71	-0.5 | 0.5	-2.0084	0.4280	2.8919	0.0870
	11	-0.4273	0.7800	04.36 | 04.79	75 | 64	-0.5 | 0.5	-0.1375	0.9740	-0.2652	0.8050
	12	-0.6207	0.6530	03.38 | 04.00	71 | 58	-0.5 | 0.5	-0.7173	0.7900	-0.5017	0.7690
	13	0.4021	0.7930	04.04 | 03.63	63 | 54	-0.5 | 0.5	0.8137	0.7630	-0.2845	0.9020
	14	1.4959	0.3120	03.72 | 02.22	58 | 50	-0.5 | 0.5	2.2768	0.3310	0.4510	0.5750
	15	-0.6667	0.6340	02.33 | 03.00	51 | 45	-0.5 | 0.5	-0.5667	0.8500	-0.2381	0.7380
	16	0.5623	0.6700	01.87 | 01.31	42 | 39	-0.5 | 0.5	1.4250	0.5200	-0.1170	0.6900
	1	-17.3024	0.0000	02.37 | 19.67	121 | 109	-0.5 | 0.5	-18.1853	0.0000	-16.2681	0.0000
	2	-1.0839	0.4080	07.32 | 08.40	121 | 109	-0.5 | 0.5	-0.1830	0.9200	-2.0728	0.2600
	3	-0.1139	0.9400	07.18 | 07.29	116 | 106	-0.5 | 0.5	-0.3569	0.8720	0.2214	0.8760
	4	-1.2606	0.3510	05.97 | 07.23	108 | 103	-0.5 | 0.5	-1.3457	0.5250	-1.0863	0.5060
	5	1.7092	0.1780	06.44 | 04.73	104 | 100	-0.5 | 0.5	1.1852	0.5670	2.2846	0.1330
	6	1.7458	0.1550	05.65 | 03.91	97 | 98	-0.5 | 0.5	1.6366	0.3630	1.9508	0.2550
	7	2.4584	0.0240	05.36 | 02.91	95 | 88	-0.5 | 0.5	2.6242	0.1090	2.3333	0.1490
Web.com Events	8	3.2425	0.0110	05.56 | 02.32	90 | 85	-0.5 | 0.5	3.9715	0.0360	2.6142	0.0870
	9	0.9159	0.4380	03.79 | 02.87	87 | 76	-0.5 | 0.5	0.4647	0.7830	1.3696	0.4110
	10	2.1414	0.0730	04.52 | 02.38	79 | 71	-0.5 | 0.5	4.9846	0.0040	-0.5946	0.7240
	11	0.2206	0.8380	03.14 | 02.92	75 | 64	-0.5 | 0.5	1.2313	0.4410	-0.8071	0.6560
	12	-0.5024	0.7100	02.95 | 03.45	71 | 58	-0.5 | 0.5	-0.7903	0.6680	-0.1684	0.9010
	13	-0.7169	0.5780	02.43 | 03.14	63 | 54	-0.5 | 0.5	-0.9980	0.5590	-0.4138	0.8160
	14	-0.8331	0.5640	02.96 | 03.79	58 | 50	-0.5 | 0.5	-1.2545	0.5620	-0.3566	0.8500
	15	0.6706	0.6230	02.87 | 02.20	51 | 45	-0.5 | 0.5	1.4750	0.5030	-0.0952	1.0000
	16	1.0895	0.4810	03.29 | 02.20	40 | 38	-0.5 | 0.5	2.6986	0.3210	-0.3216	0.7420
	1	3.7471	0.0000	25.45 | 21.70	121 | 109	-0.5 | 0.5	4.0491	0.0000	3.4290	0.0020
	2	1.8279	0.2370	18.43 | 16.60	121 | 109	-0.5 | 0.5	0.8437	0.7110	2.8299	0.1410
	3	0.9006	0.5660	17.05 | 16.15	116 | 106	-0.5 | 0.5	0.3264	0.9070	1.5286	0.5040
	4	0.4691	0.7850	14.93 | 14.46	108 | 103	-0.5 | 0.5	-0.4293	0.8490	1.6086	0.4890
	5	1.3858	0.4290	14.28 | 12.89	104 | 100	-0.5 | 0.5	-0.4954	0.8600	3.5892	0.1130
	6	3.4269	0.0590	14.59 | 11.16	97 | 98	-0.5 | 0.5	2.1285	0.4120	5.1235	0.0290
	7	3.1301	0.0790	12.91 | 09.78	95 | 88	-0.5 | 0.5	2.4958	0.3340	4.2000	0.0550
US Events	8	4.4412	0.0130	12.34 | 07.90	90 | 85	-0.5 | 0.5	4.3477	0.1240	5.0349	0.0300
	9	1.2160	0.5290	09.66 | 08.44	86 | 76	-0.5 | 0.5	0.7297	0.7910	2.3936	0.3160
	10	2.1979	0.2400	09.63 | 07.44	78 | 71	-0.5 | 0.5	2.9062	0.2750	2.1764	0.3570
	11	-0.3285	0.8440	07.47 | 07.80	74 | 64	-0.5 | 0.5	1.0563	0.7210	-1.1967	0.5870
	12	-1.2517	0.5150	06.39 | 07.64	69 | 57	-0.5 | 0.5	-1.5781	0.5750	-0.7543	0.7310
	13	-0.5695	0.7660	06.55 | 07.12	60 | 53	-0.5 | 0.5	-0.2510	0.9280	-0.9766	0.6590
	14	0.3298	0.8830	06.90 | 06.57	53 | 48	-0.5 | 0.5	0.9152	0.7270	-0.2310	0.9430
	15	-0.1984	0.9140	05.42 | 05.62	47 | 43	-0.5 | 0.5	0.9000	0.7740	-0.5913	0.6940
	16	1.9333	0.3160	05.60 | 03.67	39 | 35	-0.5 | 0.5	4.0417	0.1810	-0.5333	0.6960

Notes: *See notes to Table 2.5.*

Table 2.12: Q School Employment Results (Local Randomization)

	All						Young		Old	
	year	τ	pval	β_l | β_r	N_l | N_r	h_l | h_r	τ	pval	τ	pval
	1	-0.0009	1.0000	0.991 | 0.992	121 | 109	-0.5 | 0.5	0.0156	1.0000	-0.0189	0.4760
	2	0.0341	0.6080	0.844 | 0.810	121 | 109	-0.5 | 0.5	0.0045	1.0000	0.0636	0.3860
	3	0.0527	0.4350	0.811 | 0.759	116 | 106	-0.5 | 0.5	-0.0242	0.8240	0.1338	0.1210
	4	0.0064	1.0000	0.738 | 0.731	108 | 103	-0.5 | 0.5	-0.0464	0.6730	0.0637	0.5610
	5	0.1250	0.0730	0.750 | 0.625	104 | 100	-0.5 | 0.5	0.0486	0.6680	0.2108	0.0450
	6	0.2293	0.0000	0.755 | 0.526	97 | 98	-0.5 | 0.5	0.2111	0.0210	0.2624	0.0120
	7	0.2115	0.0000	0.727 | 0.516	95 | 88	-0.5 | 0.5	0.2107	0.0370	0.2222	0.0560
+ World Earnings	8	0.2471	0.0020	0.647 | 0.400	90 | 85	-0.5 | 0.5	0.2455	0.0290	0.2664	0.0180
	9	0.1851	0.0300	0.592 | 0.407	86 | 76	-0.5 | 0.5	0.2350	0.0340	0.1609	0.1760
	10	0.0711	0.4070	0.507 | 0.436	78 | 71	-0.5 | 0.5	0.1821	0.1590	-0.0090	1.0000
	11	0.0030	1.0000	0.422 | 0.419	74 | 64	-0.5 | 0.5	0.1000	0.4770	-0.0754	0.5930
	12	-0.0694	0.4910	0.351 | 0.420	69 | 57	-0.5 | 0.5	-0.1010	0.4680	-0.0273	1.0000
	13	-0.0204	0.8550	0.396 | 0.417	60 | 53	-0.5 | 0.5	-0.0627	0.7910	0.0351	1.0000
	14	0.0562	0.6680	0.396 | 0.340	53 | 48	-0.5 | 0.5	0.0938	0.5670	0.0119	1.0000
	15	-0.0807	0.5120	0.302 | 0.383	47 | 43	-0.5 | 0.5	0.0250	1.0000	-0.1889	0.2340
	16	0.0637	0.5960	0.371 | 0.308	39 | 35	-0.5 | 0.5	0.2583	0.1220	-0.2000	0.3720
	1	0.7172	0.0000	0.982 | 0.264	121 | 109	-0.5 | 0.5	0.8259	0.0000	0.5952	0.0000
	2	0.1097	0.1170	0.523 | 0.413	121 | 109	-0.5 | 0.5	0.0313	0.8420	0.1880	0.0440
	3	0.0595	0.4120	0.491 | 0.431	116 | 106	-0.5 | 0.5	0.0818	0.4180	0.0303	0.8440
	4	0.0965	0.1700	0.485 | 0.389	108 | 103	-0.5 | 0.5	0.0454	0.6870	0.1463	0.1650
	5	0.1050	0.1510	0.480 | 0.375	104 | 100	-0.5 | 0.5	-0.0046	1.0000	0.2200	0.0270
	6	0.0986	0.1880	0.439 | 0.340	97 | 98	-0.5 | 0.5	0.0494	0.6980	0.1581	0.1220
	7	0.0697	0.3620	0.375 | 0.305	95 | 88	-0.5 | 0.5	-0.0214	1.0000	0.1778	0.0860
+ PGA TOUR Earnings	8	0.0510	0.4950	0.318 | 0.267	90 | 85	-0.5 | 0.5	0.0104	1.0000	0.1110	0.2460
	9	0.0220	0.8570	0.289 | 0.267	86 | 76	-0.5 | 0.5	0.0194	1.0000	0.0538	0.5350
	10	-0.0144	0.8520	0.268 | 0.282	78 | 71	-0.5 | 0.5	-0.0756	0.6510	0.0781	0.5160
	11	-0.0068	1.0000	0.250 | 0.257	74 | 64	-0.5 | 0.5	0.0313	0.7890	-0.0239	1.0000
	12	-0.0038	1.0000	0.228 | 0.232	69 | 57	-0.5 | 0.5	0.0068	1.0000	-0.0136	1.0000
	13	0.0453	0.6310	0.245 | 0.200	60 | 53	-0.5 | 0.5	0.0392	0.7830	0.0535	0.6550
	14	0.0594	0.6300	0.229 | 0.170	53 | 48	-0.5 | 0.5	0.0714	0.5760	0.0524	0.6090
	15	-0.0307	0.7810	0.140 | 0.170	47 | 43	-0.5 | 0.5	0.0167	1.0000	-0.0588	0.4820
	16	-0.0652	0.5050	0.114 | 0.179	39 | 35	-0.5 | 0.5	-0.0083	1.0000	-0.1333	0.5030
	1	-0.4137	0.0000	0.495 | 0.909	121 | 109	-0.5 | 0.5	-0.2634	0.0000	-0.5727	0.0000
	2	-0.0482	0.5190	0.514 | 0.562	121 | 109	-0.5 | 0.5	-0.0179	0.8690	-0.0844	0.4010
	3	0.0407	0.5960	0.472 | 0.431	116 | 106	-0.5 | 0.5	-0.0305	0.8500	0.1156	0.2660
	4	-0.0168	0.8950	0.437 | 0.454	108 | 103	-0.5 | 0.5	-0.0485	0.6930	0.0098	1.0000
	5	0.0738	0.3430	0.420 | 0.346	104 | 100	-0.5 | 0.5	0.0833	0.4200	0.0631	0.5330
	6	0.1400	0.0470	0.418 | 0.278	97 | 98	-0.5 | 0.5	0.1894	0.0590	0.0976	0.4000
	7	0.1118	0.1090	0.375 | 0.263	95 | 88	-0.5 | 0.5	0.1786	0.0740	0.0444	0.8400
+ Web.com Earnings	8	0.1516	0.0250	0.329 | 0.178	90 | 85	-0.5 | 0.5	0.2200	0.0320	0.0867	0.4340
	9	0.1019	0.1200	0.276 | 0.174	86 | 76	-0.5 | 0.5	0.1234	0.2180	0.0814	0.4520
	10	0.0894	0.2560	0.282 | 0.192	78 | 71	-0.5 | 0.5	0.2157	0.0450	-0.0330	0.7670
	11	0.0608	0.3990	0.250 | 0.189	74 | 64	-0.5 | 0.5	0.1438	0.2080	-0.0202	1.0000
	12	-0.0359	0.6750	0.211 | 0.246	69 | 57	-0.5 | 0.5	-0.0840	0.5950	0.0248	1.0000
	13	-0.0591	0.5260	0.208 | 0.267	60 | 53	-0.5 | 0.5	-0.1235	0.3770	0.0251	1.0000
	14	0.0008	1.0000	0.208 | 0.208	53 | 48	-0.5 | 0.5	-0.0670	0.7690	0.1048	0.4260
	15	-0.0267	0.8230	0.186 | 0.213	47 | 43	-0.5 | 0.5	0.0167	1.0000	-0.0712	0.6460
	16	0.0462	0.7430	0.200 | 0.154	39 | 35	-0.5 | 0.5	0.0417	1.0000	0.0667	1.0000

Notes: *See notes to Table 2.5.*

Table 2.13: Q School Performance Results (Local Randomization)

	All						Young		Old	
	year	τ	pval	$\beta_l \mid \beta_r$	$N_l \mid N_r$	$h_l \mid h_r$	τ	pval	τ	pval
Scoring Average	1	-0.1505	0.2260	0.203 | 0.353	119 | 109	-0.5 | 0.5	-0.0314	0.8370	-0.2751	0.1450
	2	0.0347	0.8410	0.269 | 0.234	101 | 93	-0.5 | 0.5	0.2132	0.4070	-0.1279	0.5290
	3	0.0242	0.9150	0.309 | 0.285	92 | 90	-0.5 | 0.5	-0.0435	0.8910	0.0845	0.7480
	4	-0.1366	0.4150	0.186 | 0.323	83 | 82	-0.5 | 0.5	-0.1565	0.5870	-0.1398	0.4960
	5	-0.3397	0.1360	0.257 | 0.597	70 | 77	-0.5 | 0.5	-0.3331	0.2400	-0.4249	0.2590
	6	-0.1575	0.5430	0.397 | 0.554	61 | 72	-0.5 | 0.5	-0.0974	0.8060	-0.3093	0.3770
	7	0.3454	0.2430	0.806 | 0.460	53 | 69	-0.5 | 0.5	0.2873	0.4300	0.2607	0.5220
	8	0.1425	0.7260	0.601 | 0.458	42 | 57	-0.5 | 0.5	-0.0788	0.8750	0.2057	0.7250
	9	-0.4352	0.3320	0.381 | 0.816	42 | 41	-0.5 | 0.5	-0.1698	0.6320	-1.1421	0.2630
	10	-0.5121	0.4310	0.756 | 1.268	39 | 37	-0.5 | 0.5	-1.1742	0.1420	0.6300	0.4470
	11	-0.1375	0.7010	0.468 | 0.605	32 | 27	-0.5 | 0.5	0.0133	0.9810	-0.5260	0.5680
	12	-0.5037	0.3500	0.490 | 0.994	33 | 21	-0.5 | 0.5	-0.4380	0.4620	-0.7637	0.5640
	13	-0.7151	0.1410	0.407 | 1.122	29 | 21	-0.5 | 0.5	-0.4328	0.3500	-1.2266	0.2700
	14	-0.2953	0.7090	0.891 | 1.186	21 | 19	-0.5 | 0.5	-1.0616	0.0640	1.5583	0.7280
	15	-0.3333	0.5490	0.320 | 0.654	18 | 13	-0.5 | 0.5	-0.5426	0.3840	0.5571	0.7320
	16	-0.4124	0.6470	1.046 | 1.459	14 | 13	-0.5 | 0.5	-0.7613	0.4760	0.6418	0.7040
Unadj SA	1	0.7136	0.0000	0.497 | -0.217	119 | 109	-0.5 | 0.5	0.8305	0.0000	0.5874	0.0010
	2	0.2055	0.1500	0.207 | 0.001	101 | 93	-0.5 | 0.5	0.3220	0.1620	0.0992	0.5670
	3	0.0661	0.7470	0.204 | 0.138	92 | 90	-0.5 | 0.5	0.0581	0.8580	0.0566	0.8480
	4	-0.0252	0.8600	0.092 | 0.117	83 | 82	-0.5 | 0.5	-0.0868	0.7130	0.0299	0.8770
	5	-0.3581	0.0750	0.147 | 0.506	70 | 77	-0.5 | 0.5	-0.4223	0.0810	-0.3426	0.2900
	6	-0.1248	0.5610	0.329 | 0.454	61 | 72	-0.5 | 0.5	-0.1618	0.6290	-0.1268	0.6840
	7	0.2512	0.3450	0.685 | 0.434	53 | 69	-0.5 | 0.5	0.0779	0.7820	0.3323	0.3900
	8	-0.0075	0.9860	0.505 | 0.512	42 | 57	-0.5 | 0.5	-0.3665	0.4030	0.3486	0.5530
	9	-0.5504	0.1770	0.268 | 0.818	42 | 41	-0.5 | 0.5	-0.2590	0.3720	-1.2400	0.2250
	10	-0.6572	0.2600	0.610 | 1.268	39 | 37	-0.5 | 0.5	-1.4676	0.0250	0.8264	0.2920
	11	-0.2018	0.5450	0.358 | 0.559	32 | 27	-0.5 | 0.5	-0.0253	0.9510	-0.6953	0.4340
	12	-0.4104	0.4070	0.404 | 0.814	33 | 21	-0.5 | 0.5	-0.2376	0.6450	-0.8969	0.4990
	13	-0.6040	0.1800	0.343 | 0.947	29 | 21	-0.5 | 0.5	-0.1785	0.6310	-1.4060	0.2300
	14	-0.1226	0.8550	0.799 | 0.922	21 | 19	-0.5 | 0.5	-0.7722	0.1430	1.5369	0.6600
	15	-0.2653	0.6020	0.203 | 0.468	18 | 13	-0.5 | 0.5	-0.4626	0.4320	0.4424	0.7970
	16	-0.2880	0.7450	0.850 | 1.138	14 | 13	-0.5 | 0.5	-0.6215	0.5210	0.7025	0.5110

Notes: *See notes to Table 2.5.*

Table 2.14: Q School Total Earnings Effects (Local Randomization)

	All						Young		Old	
	year	τ	pval	$\beta_l \mid \beta_r$	$N_l \mid N_r$	$h_l \mid h_r$	τ	pval	τ	pval
World Earnings	1	265,958	0.0000	489,360 | 223,403	121 | 109	-0.50 | 0.50	224,709	0.0000	307,229	0.0090
	2	68,053	0.4300	428,932 | 360,879	121 | 109	-0.50 | 0.50	24,919	0.8440	116,623	0.2760
	3	119,505	0.1500	424,069 | 304,564	116 | 106	-0.50 | 0.50	114,910	0.3970	129,624	0.1690
	4	35,226	0.7270	386,944 | 351,718	108 | 103	-0.50 | 0.50	107,503	0.5630	-21,888	0.7740
	5	131,451	0.2150	423,911 | 292,460	104 | 100	-0.50 | 0.50	205,990	0.2290	77,712	0.4440
	6	211,853	0.0390	451,358 | 239,505	97 | 98	-0.50 | 0.50	203,640	0.2480	236,984	0.0350
	7	56,721	0.5900	380,612 | 323,891	95 | 88	-0.50 | 0.50	43,531	0.8520	102,026	0.1180
	8	184,863	0.1030	397,265 | 212,402	90 | 85	-0.50 | 0.50	313,065	0.1480	86,568	0.2310
	9	45,238	0.7320	354,217 | 308,979	86 | 76	-0.50 | 0.50	98,106	0.6850	36,506	0.5750
	10	85,447	0.5410	274,899 | 189,452	78 | 71	-0.50 | 0.50	205,880	0.4050	153	0.9980
	11	-9,731	0.9210	173,919 | 183,650	74 | 64	-0.50 | 0.50	63,271	0.6960	-67,926	0.1390
	12	117,797	0.4060	273,067 | 155,270	69 | 57	-0.50 | 0.50	265,541	0.2770	-56,038	0.1550
	13	174,616	0.3610	293,375 | 118,760	60 | 53	-0.50 | 0.50	334,356	0.2800	-33,546	0.3950
	14	95,109	0.5170	200,142 | 105,033	53 | 48	-0.50 | 0.50	188,650	0.4250	-31,238	0.5800
	15	90,785	0.5410	190,375 | 099,590	47 | 43	-0.50 | 0.50	188,788	0.4150	-8,853	0.3000
	16	105,329	0.5590	201,931 | 096,602	39 | 35	-0.50 | 0.50	204,203	0.5500	-12,644	0.0600
PGA TOUR Earnings	1	409,883	0.0000	461,519 | 051,636	121 | 109	-0.50 | 0.50	366,195	0.0000	453,083	0.0000
	2	107,395	0.1970	363,269 | 255,874	121 | 109	-0.50 | 0.50	64,184	0.6340	155,467	0.1190
	3	129,320	0.1180	369,423 | 240,103	116 | 106	-0.50 | 0.50	113,308	0.4190	149,922	0.1080
	4	71,633	0.4560	333,409 | 261,776	108 | 103	-0.50 | 0.50	132,042	0.4810	25,224	0.7070
	5	113,084	0.2840	354,164 | 241,079	104 | 100	-0.50 | 0.50	162,823	0.3680	81,550	0.4070
	6	195,325	0.0520	397,616 | 202,291	97 | 98	-0.50 | 0.50	172,544	0.3290	233,010	0.0410
	7	53,628	0.6000	331,567 | 277,940	95 | 88	-0.50 | 0.50	26,033	0.9090	112,161	0.0530
	8	190,079	0.0910	363,689 | 173,610	90 | 85	-0.50 | 0.50	316,053	0.1400	91,985	0.1880
	9	62,113	0.6520	331,177 | 269,064	86 | 76	-0.50 | 0.50	111,923	0.6500	56,624	0.3460
	10	86,500	0.5480	242,599 | 156,099	78 | 71	-0.50 | 0.50	177,322	0.4890	31,280	0.2490
	11	857	0.9930	153,994 | 153,137	74 | 64	-0.50 | 0.50	44,748	0.7750	-25,321	0.8720
	12	126,172	0.3710	252,639 | 126,467	69 | 57	-0.50 | 0.50	261,226	0.2780	-32,480	0.6630
	13	176,603	0.3510	283,478 | 106,876	60 | 53	-0.50 | 0.50	331,842	0.2790	-25,686	0.9810
	14	105,081	0.4710	190,286 | 085,205	53 | 48	-0.50 | 0.50	199,726	0.3850	-23,410	1.0000
	15	91,514	0.5430	178,677 | 087,163	47 | 43	-0.50 | 0.50	184,996	0.4250	-2,510	0.4820
	16	104,022	0.5640	192,438 | 088,416	39 | 35	-0.50 | 0.50	199,002	0.5610	-9,458	0.5030
Web.com Earnings	1	-111,883	0.0000	012,266 | 124,149	121 | 109	-0.50 | 0.50	-125,255	0.0000	-96,670	0.0000
	2	-23,050	0.0740	035,802 | 058,852	121 | 109	-0.50 | 0.50	-14,570	0.4910	-31,982	0.0370
	3	-9,962	0.3570	032,580 | 042,542	116 | 106	-0.50 | 0.50	-6,813	0.7070	-12,407	0.2240
	4	-13,539	0.3050	035,293 | 048,832	108 | 103	-0.50 | 0.50	-10,980	0.5890	-14,754	0.2900
	5	12,518	0.2450	039,207 | 026,689	104 | 100	-0.50 | 0.50	13,263	0.4620	12,699	0.3700
	6	5,376	0.5120	022,428 | 017,051	97 | 98	-0.50 | 0.50	-1,152	0.9360	12,600	0.1070
	7	8,695	0.1420	018,375 | 009,680	95 | 88	-0.50 | 0.50	15,616	0.1110	2,145	0.7380
	8	7,224	0.3810	022,573 | 015,349	90 | 85	-0.50 | 0.50	12,197	0.3790	3,490	0.5910
	9	3,398	0.6570	016,254 | 012,857	86 | 76	-0.50 | 0.50	2,579	0.8160	4,284	0.6320
	10	11,044	0.1890	022,495 | 011,452	78 | 71	-0.50 | 0.50	20,281	0.1840	3,228	0.6970
	11	-422	0.9610	010,407 | 010,828	74 | 64	-0.50 | 0.50	7,539	0.5140	-8,766	0.3660
	12	-406	0.9640	013,571 | 013,976	69 | 57	-0.50 | 0.50	3,263	0.8190	-4,666	0.4420
	13	736	0.8870	008,576 | 007,839	60 | 53	-0.50 | 0.50	3,355	0.7040	-2,679	0.8960
	14	-4,630	0.5540	009,558 | 014,188	53 | 48	-0.50 | 0.50	-7,862	0.5470	730	0.8500
	15	2,278	0.7480	011,669 | 009,391	47 | 43	-0.50 | 0.50	5,017	0.6280	-204	0.9660
	16	2,067	0.7120	009,202 | 007,135	39 | 35	-0.50 | 0.50	4,692	0.6390	-453	1.0000

Notes: *These regressions are estimated in earnings levels rather than logs. See notes to Table 2.5.*

2.11 Appendix: Robustness

2.11.1 Robustness across Specifications

The following figures display results of the ATE's for each experiment under different specifications. In each figure the legend provides detailed information in each specification. "BW" denotes bandwidth selection method with the options of "CER" (coverage error optimal, see Cattaneo et al. (2016c)) or "MSE" (minimum mean-squared error). "PD" denote polynomial degree and varies between a first or second order polynomial. "TE" refers to the method of constructing the ATE and can either be "BC" (biased-corrected as recommended by Calonico et al. (2014)) or "Conv" (conventional). When employing the bias-corrected treatment effect method I set the "SE" (standard error) method to "Robust" which denotes a method robust to specification error as detailed in Calonico et al. (2014). When employing the conventional treatment effect method I set the standard error method to "Conv" (convential) to denote standard heteroskedasitic standard errors.

Figure 2.32 shows the results of varying the bandwidth selection method and polynomial degree. Figure 2.33 shows the results of varying the estimation procedure between the robust, bias-corrected method of Calonico et al. (2014) versus more conventional methods. In Figures 2.32 and 2.33 both sets of results use local linear methods (as opposed to the local randomization method). Figure 2.34 compares the results of local randomization method and the local linear method in the main specification (CER bandwidth, linear, bias-corrected, robust). Figures 2.35 through 2.37 repeat these exercises for the results for future performance. Figures 2.38 through 2.40 show the results for future events played. Figures 2.41 through 2.43 show the results for future employment.

Figure 2.32: Local Polynomial Specification Robustness of ATE's on Future World Earnings

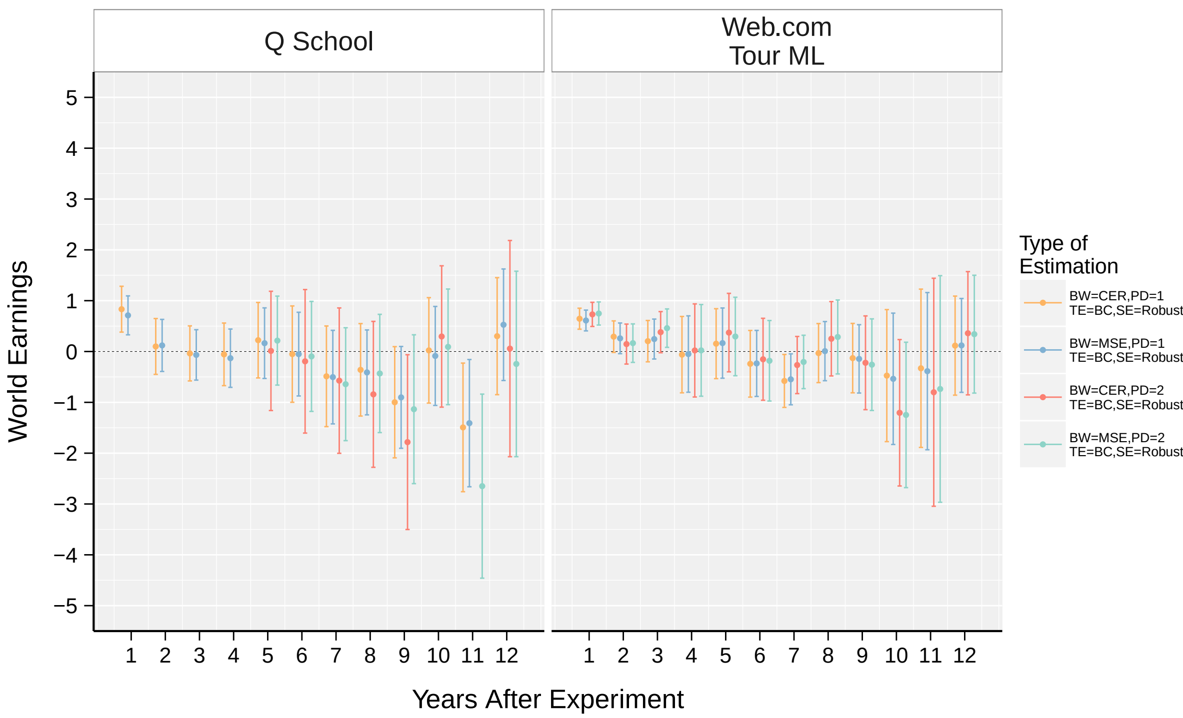


Figure 2.33: Bias-Correction Robustness of ATE's on Future World Earnings

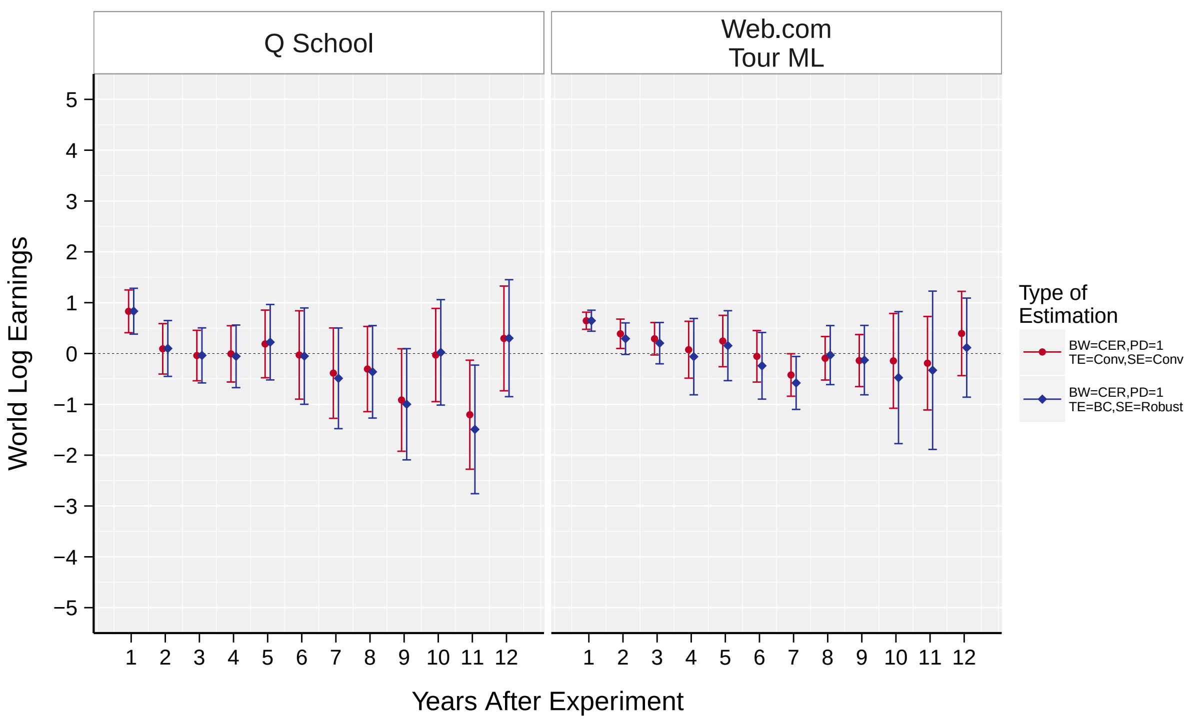


Figure 2.34: Discrete vs. Continuous Running Variable Method Robustness of ATE's on Future World Earnings

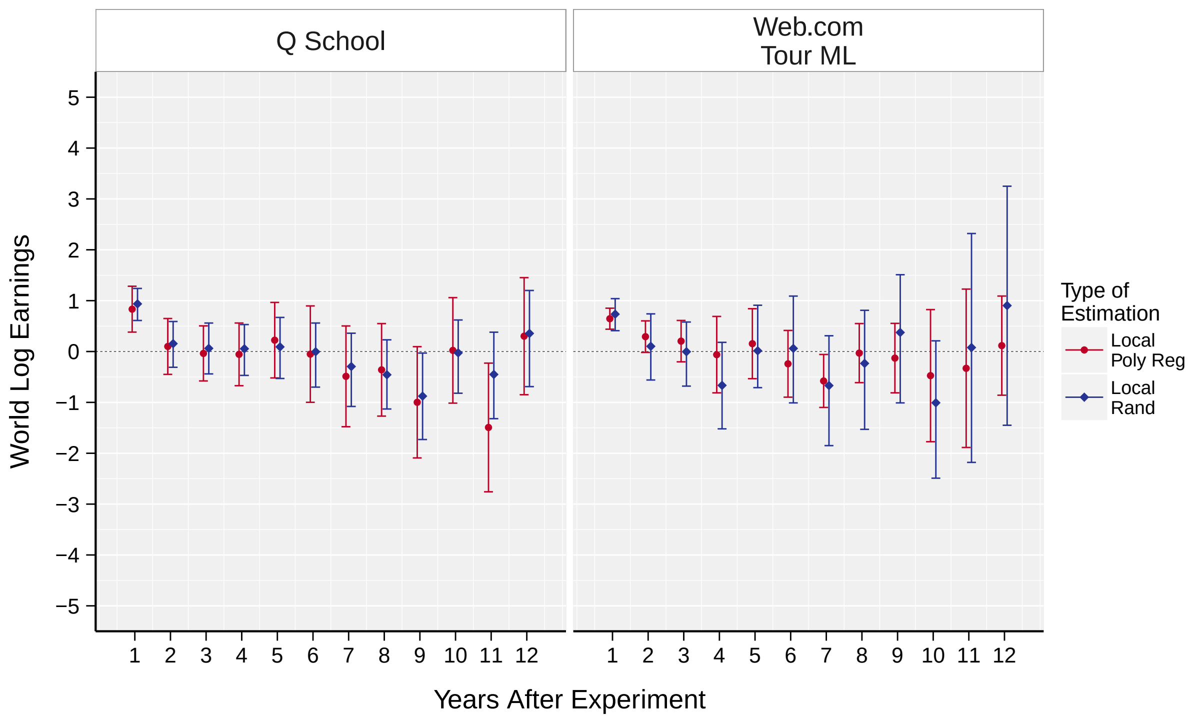


Figure 2.35: Local Polynomial Specification Robustness of ATE's on Future Scoring Average

Figure 2.36: Bias-Correction Robustness of ATE's on Future Scoring Average

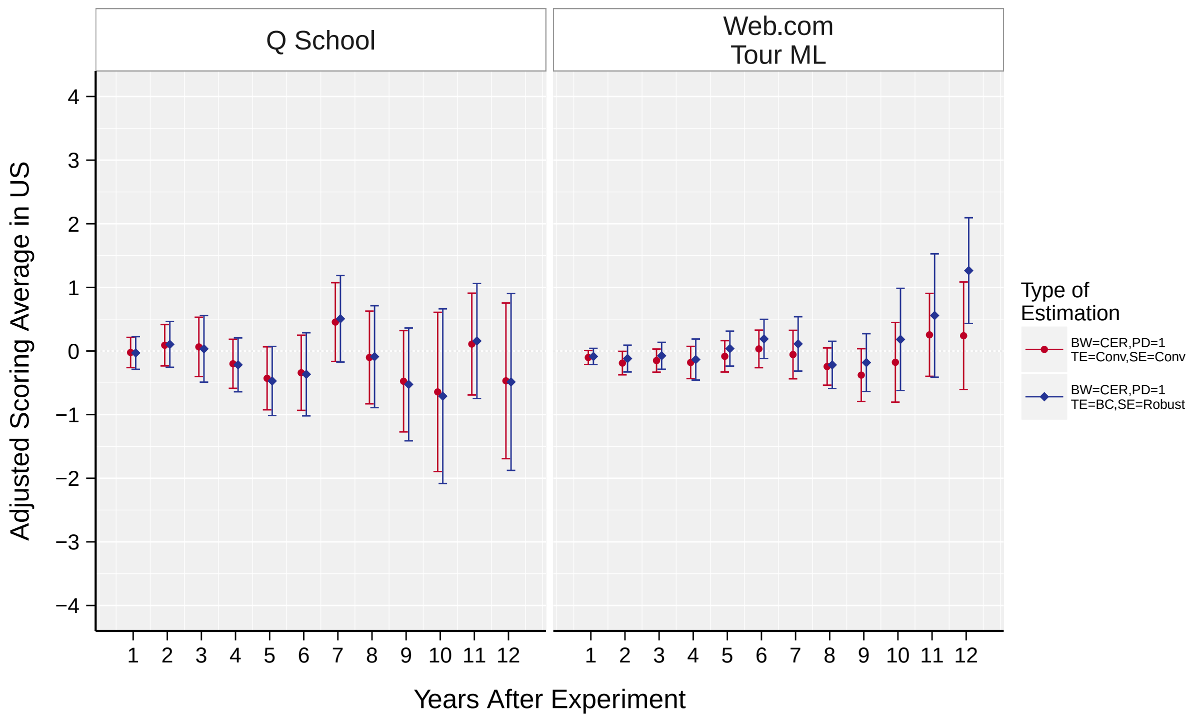


Figure 2.37: Discrete vs. Continuous Running Variable Method Robustness of ATE's on Future Scoring Average

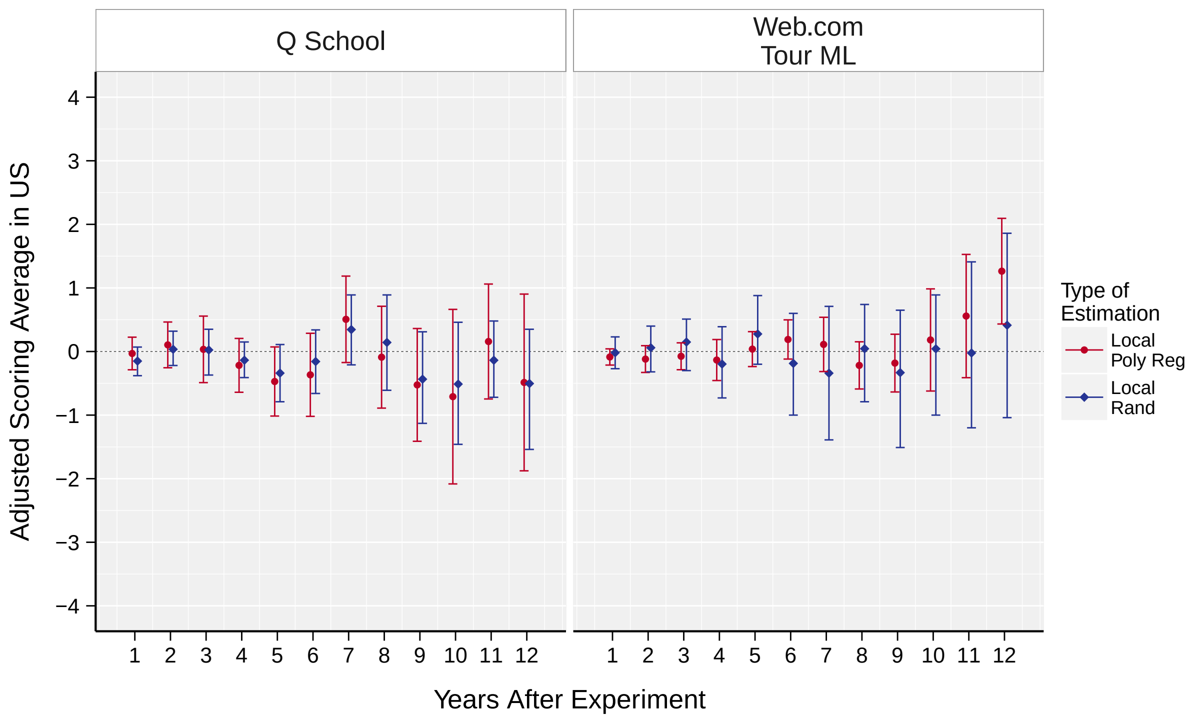


Figure 2.38: Local Polynomial Specification Robustness of ATE's on Future PGA TOUR Events

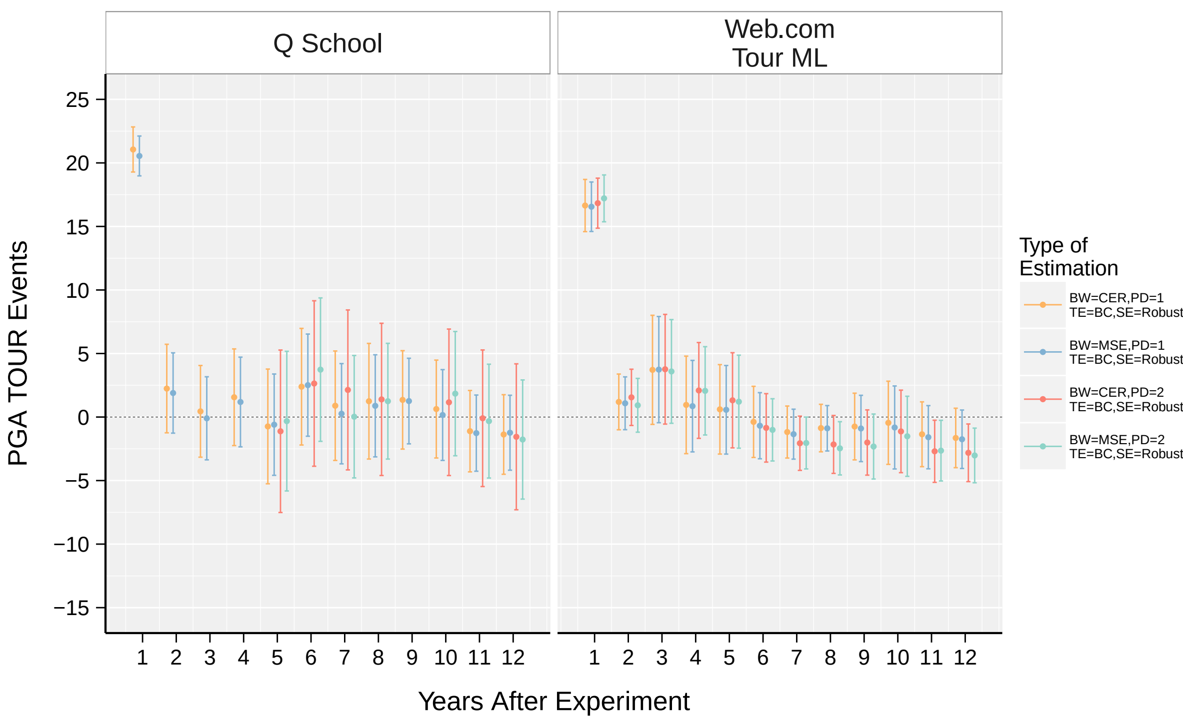


Figure 2.39: Bias-Correction Robustness of ATE's on Future PGA TOUR Events

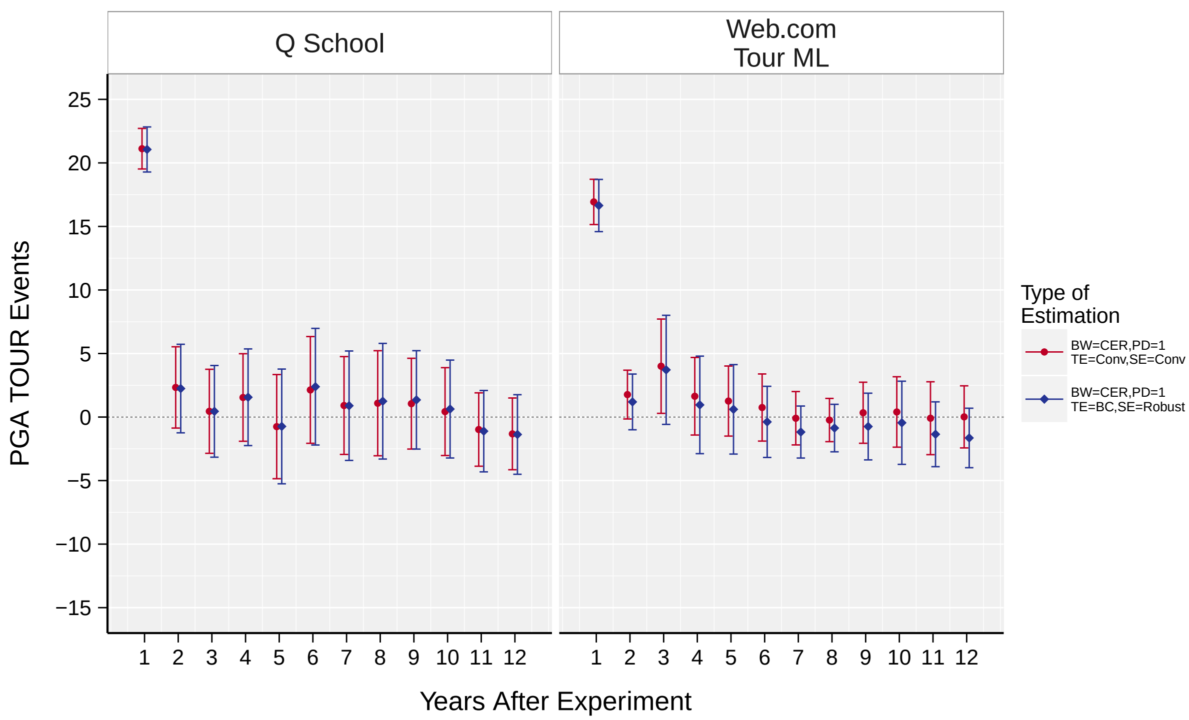


Figure 2.40: Discrete vs. Continuous Running Variable Method Robustness of ATE's on Future PGA TOUR Events

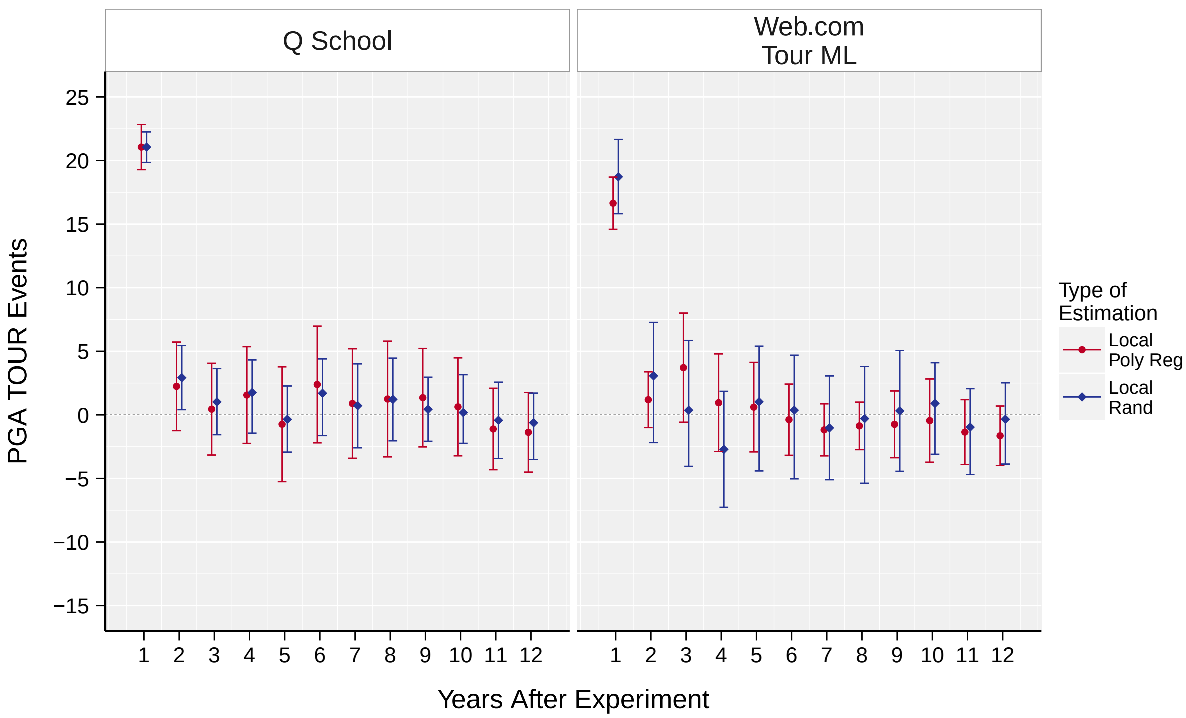


Figure 2.41: Local Polynomial Specification Robustness of ATE's on Future Positive World Earnings

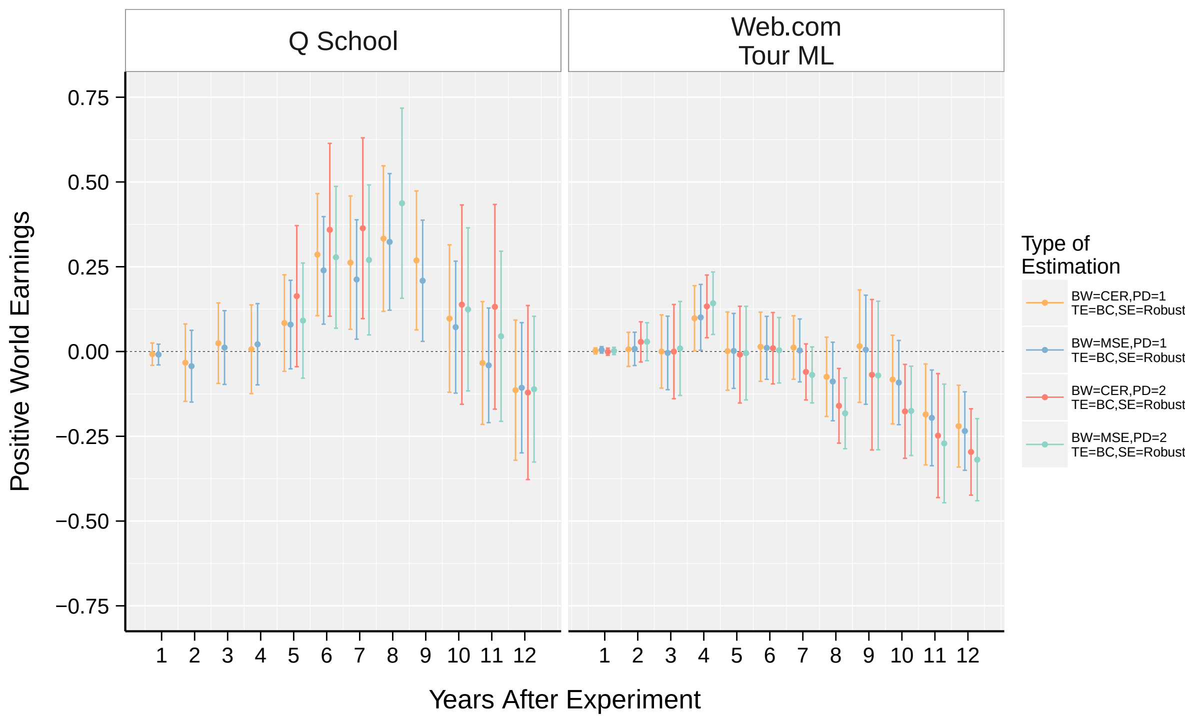


Figure 2.42: Bias-Correction Robustness of ATE's on Future Positive World Earnings

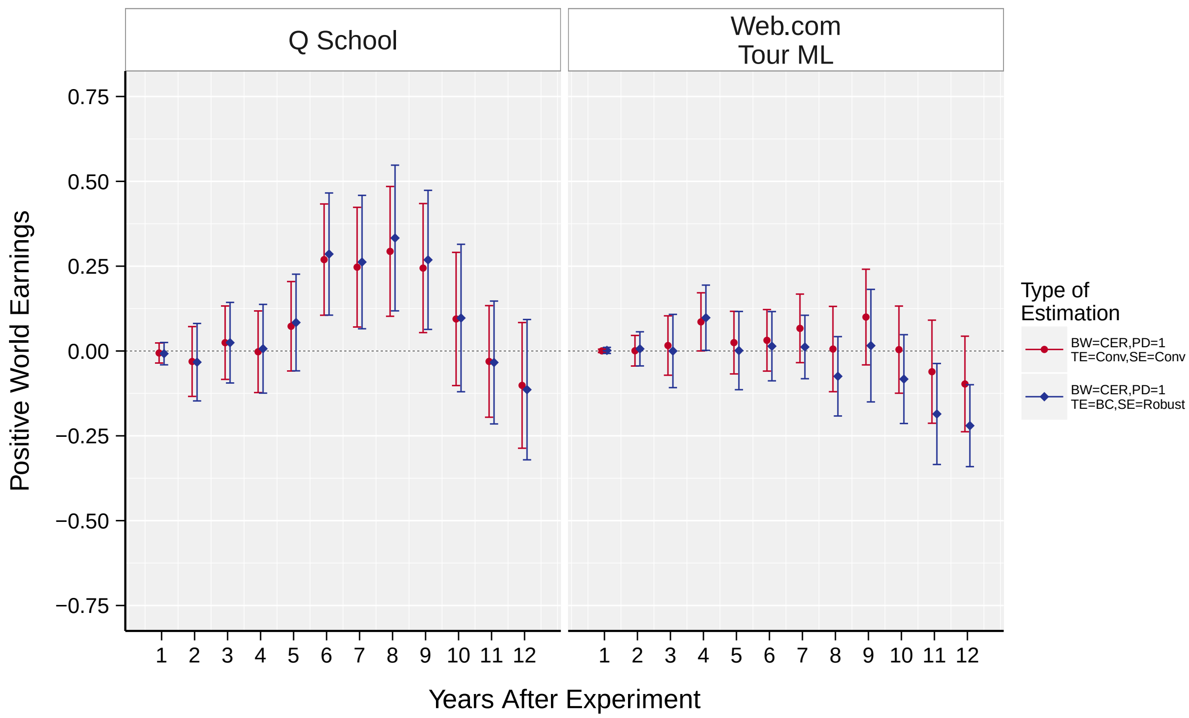
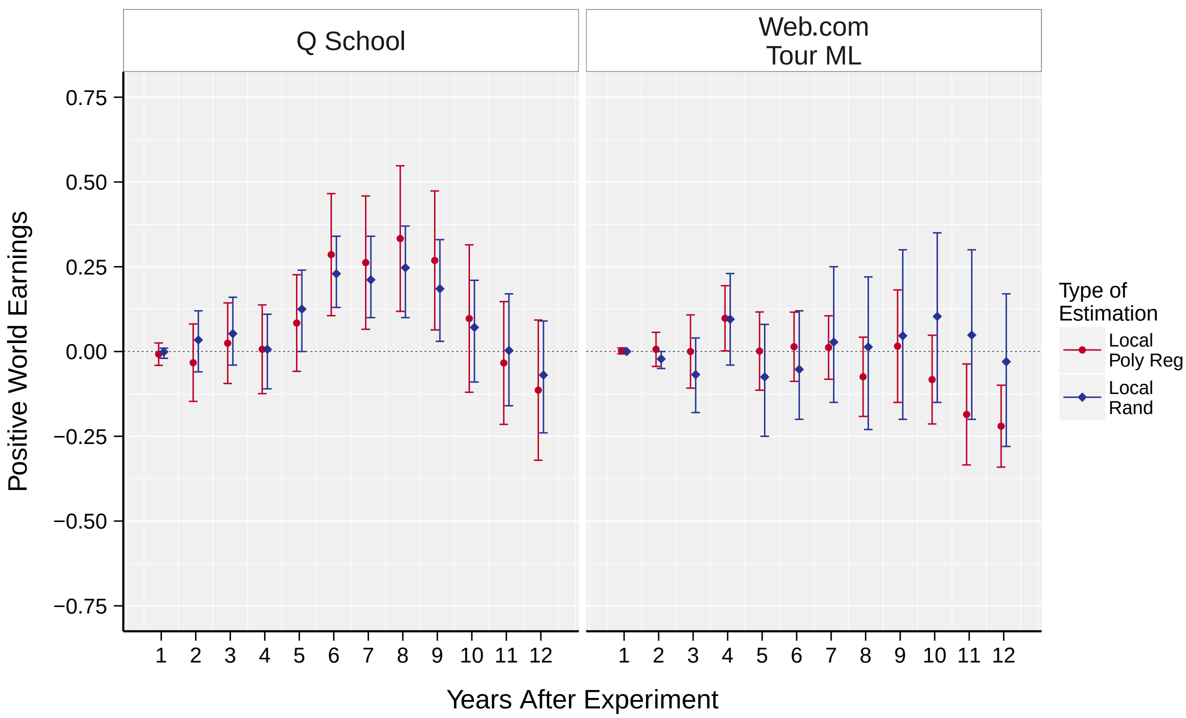


Figure 2.43: Discrete vs. Continuous Running Variable Method Robustness of ATE's on Future Positive World Earnings



2.11.2 Robustness over Time

The following tables report the estimated ATE's for both experiments for an early and a later period, in each case using the main specification. Table 2.15 presents very similar effects on log world earnings in the 1990's versus the 2000's for the Web.com Tour ML experiment. Table 2.16 shows generally similar results across time period for the Q School experiment. The pattern of a significant treatment effect in the first year followed by insignificant effects in later years holds. However, there is some evidence that the treatment effect grew strong over time as the estimated effect is about 50% larger in the 2000's versus the 1990's.

Table 2.15: Web.com Tour ML Earnings Effects over Time (Local Linear)

	year	τ	se	pval	$\beta_l \mid \beta_r$	$N_l \mid N_r$	$b_l \mid b_r$
	1	0.4885	0.1756	0.0056	13.12 | 12.57	104 | 442	-22.02 | 46.50
World Earnings	2	0.3644	0.2496	0.1449	12.71 | 12.30	102 | 496	-20.79 | 62.18
1990-1999	3	-0.4516	0.2778	0.1047	11.92 | 12.33	100 | 391	-22.39 | 51.10
	4	-0.4176	0.3278	0.2033	11.92 | 12.30	97 | 386	-20.49 | 57.06
	5	-0.4353	0.2732	0.1120	11.93 | 12.29	87 | 267	-22.06 | 40.04
	1	0.5032	0.1130	0.0000	12.71 | 12.20	193 | 523	-21.85 | 53.59
World Earnings	2	-0.2209	0.1527	0.1485	11.98 | 12.22	182 | 552	-20.48 | 65.53
2000-2009	3	0.1952	0.3162	0.5372	12.19 | 12.01	185 | 565	-23.16 | 76.33
	4	-0.1546	0.4738	0.7443	11.76 | 11.92	182 | 410	-23.71 | 57.17
	5	0.2439	0.5589	0.6627	11.92 | 11.74	176 | 364	-27.27 | 57.10

Table 2.16: Q School Earnings Effects over Time (Local Randomization)

	year	τ	pval	$\beta_l \mid \beta_r$	$N_l \mid N_r$	$h_l \mid b_r$
	1	0.7359	0.0030	12.44 | 11.70	62 | 53	-0.50 | 0.50
World Earnings	2	0.0931	0.8000	12.22 | 12.13	47 | 44	-0.50 | 0.50
1993-2001	3	0.2421	0.5120	12.11 | 11.87	45 | 44	-0.50 | 0.50
	4	0.1769	0.6190	12.33 | 12.16	45 | 39	-0.50 | 0.50
	5	0.1687	0.6780	12.16 | 11.99	39 | 41	-0.50 | 0.50
	1	1.2351	0.0000	12.71 | 11.48	41 | 46	-0.50 | 0.50
World Earnings	2	0.0391	0.9140	12.07 | 12.04	35 | 42	-0.50 | 0.50
2002-2009	3	-0.5665	0.1300	11.87 | 12.44	32 | 38	-0.50 | 0.50
	4	-0.1691	0.6580	11.70 | 11.87	31 | 36	-0.50 | 0.50
	5	0.0136	0.9740	11.87 | 11.86	26 | 34	-0.50 | 0.50

Chapter 3

The Disappearing Large-Firm Wage Premium

with Nicholas Bloom, Fatih Guvenen, Jae Song, and Till von Wachter

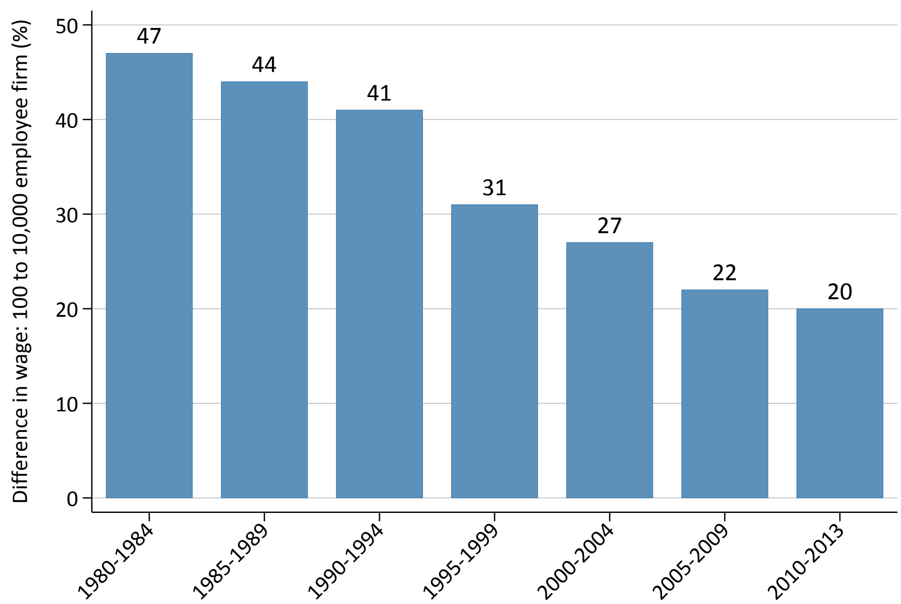
3.1 Introduction

Large firms pay higher wages than smaller firms even after controlling for the quality of a worker. While this empirical fact may seem surprising, it has been shown to hold true in many countries during most of the 20th century, going back to the first analysis of the subject by Moore (1911).[1] In this paper, we show that the large-firm wage premium (LFWP)—which we define as the gap between the average wage earnings of employees in large versus small firms (without controlling for worker characteristics)—has declined significantly since the early 1980s. The simplest illustration of this fact—and our first main result—is shown in Figure 3.1, which plots the LFWP in each

[1] Among many others, see Brown and Medoff (1989). Oi and Idson (1999) provide a thorough survey of the vast literature on the subject.

5-year period between a 100-employee firm (about the 25th percentile of employment-weighted firm-size distribution) to a 10,000-employee firm (about the 75th percentile). We see that the LFWP declined from 47% (i.e, average worker earnings in the latter firm size category was 47% higher than its counterpart in the former category) in the early 1980s to 20% by the early 2010s.

Figure 3.1: Difference in Average Worker Earnings: a 10,000-Employee Firm Minus a 100-Employee Firm



In the rest of the paper, we expand on this main result in several important directions. First, any discussion of the LFWP that does not address selection seriously would be incomplete. We will address two sides of the selection problem: (i) workers employed by large and small firms are likely to be different along various dimensions (e.g., skill or education, age, gender, etc.), and (ii) large and small firms themselves likely differ in various ways—in terms of industry composition, geographical regions they operate in, and so on. These points are well understood in the extant literature on the *level* of the premium, and has been addressed by using various controls.[2] Of

[2]Moore (1911) was keenly aware of this issue. His analysis was based on female workers in Italian textile factories

course, similar selection issues could also be contributing to the *decline* in the LFWP *over time*.

To explore this point, we conduct several exercises. We use the fixed-effects regression framework of Abowd et al. (1999) [AKM], which allows us to estimate a separate wage fixed effect for each worker and for each firm in our sample. The worker fixed effect (or "worker quality") can be thought of as capturing both observable *and* unobservable characteristics that allow the worker to earn a high wage (controlling for the employer pay premium). The firm fixed effect is interpreted similarly as the premium a firm pays to a typical worker (i.e., quality-adjusted) relative to what the average firm pays.

During our sample period, we find that the average worker fixed effect in a firm rises with the size of the firm and can explain about 20% of the LFWP, which is consistent with a broad range of previous evidence that large firms hire higher-quality workers. A more novel finding of our paper is that about 70% of the LFWP can be explained by the fact that the firm pay premium strongly increases with firm size. Turning to the change over time, we show that the reduction in the LFWP stemmed from a decline in the pay premium (or firm fixed effect) large firms were paying relative to smaller firms over this period. In contrast, we find average worker quality at larger relative to smaller firms has remained stable over time.

A third result we establish is that the decline in the LFWP was concentrated primarily among very large firms—the previously substantial LFWP between a 1,000-employee firm and a 10,000-employee firm has effectively disappeared, while the premium between a 100-employee firm and a 1,000-employee firm has only declined modestly. Again, the differential effect is almost entirely explained by a reduction in the firm pay premiums (or firm fixed effects), indicating that explanations of this decline must recognize the evolution of very large employers.

A final finding is that the bulk of the fall in the LFWP took place within industries. Although industries with a historically high LFWPs (e.g., manufacturing) shrunk while those with smaller

---

categorized by age groups, which was the most detail available at the time.

size premiums (e.g., services and retail) expanded, the shift in industry composition can only account for about 20% of the overall decline in the premium leaving 80% of the decline within industries.

The prior literature offers several potential explanations for how firm pay premiums (i.e., firm fixed effects) could rise with firm size. One hypothesis underlying the LFWP is that larger firms may be more unpleasant to work in and hence pay compensating differentials. However, as Katz and Summers (1989) show, larger firms have a far higher number of applicants per vacancy and lower quit rates, suggesting the jobs are more desirable in general. Larger firms also have higher work-life balance and other employee satisfaction metrics (Bloom et al. 2011). This suggests that compensating differentials cannot account for the majority of the LFWP. A common alternative hypothesis has been that larger firms may earn higher rents and share some of these rents with their workers. The sharing of rents could be because of perceptions of fairness, because workers and firms bargain over the surplus, or because of the presence or threat of unions. Another explanation is that larger firms may face particular challenges in monitoring their workers, and hence pay higher wages to solve personnel problems. For example, it has long been hypothesized that large firms pay efficiency wages (e.g., Krueger and Summers (1988)).

Studying the LFWP has importance beyond understanding the determinants of earnings. Large firms have been traditionally a source of high-quality jobs, especially for low-skilled workers. Changes in the availability of such jobs lead to a reduction in earnings among lower-skilled individuals, potentially raising the outlays for government programs that effectively insure low-income workers against lifetime earnings reductions, such as Old Age Survivor Insurance and also Social Security Disability Insurance. Since high-wage jobs are also those that are safer, have higher benefits, and have better working conditions (Maestas et al. (2017)), changes in how large firms treat workers could also impact the rate of claiming for these programs.

3.2 Data

We use data from the Master Earnings File (MEF), which is a confidential database compiled and maintained by the U.S. Social Security Administration (SSA). The MEF contains a separate line of record for every individual that has ever been issued a U.S. Social Security number. In addition to basic demographic information (sex, date of birth, etc.), the MEF contains labor earnings information for every year from 1978 to (as of this writing) 2013. Earnings data in the MEF are based on Box 1 of Form W-2, which is sent directly from employers to the SSA. Data from Box 1 are uncapped and include wages and salaries, bonuses, tips, exercised stock options, the dollar value of vested restricted stock units, and other sources of income deemed as remuneration for labor services by the U.S. Internal Revenue Service. The SSA MEF data are described in detail in Olsen and Hudson (2009).

3.3 Econometric Model

To analyze worker and firm components of earnings we follow the Card et al. (2013) implementation of the model introduced by Abowd et al. (1999). We will divide our time period into five seven-year periods and estimate a separate model for each period p. The regression model we estimate in each period is

$$y_t^{i,j} = \theta^{i,p} + X_t^i \beta^p + \psi^{j,p} + \epsilon_t^{i,j}, \tag{3.1}$$

where $\theta^{i,p}$ is the worker fixed effect which captures earnings differences due to fixed worker characteristics (such as returns to schooling or to innate ability) that are unobservable by the econometrician , β^p is a vector that captures the effects of time-varying and observable worker characteristics and aggregate shocks (in our case, a polynomial in age and year effects), and $\psi^{j,p}$ is the firm fixed effect which captures the wage premium firm j pays relative to other firms for the same quality-adjusted worker (which may be due to rent sharing or compensating differentials). The residual, $\epsilon_t^{i,j}$, captures transitory earnings fluctuations. We leave the dependence of the identity of the firm on the worker implicit, such that $j \equiv j(i,t)$.

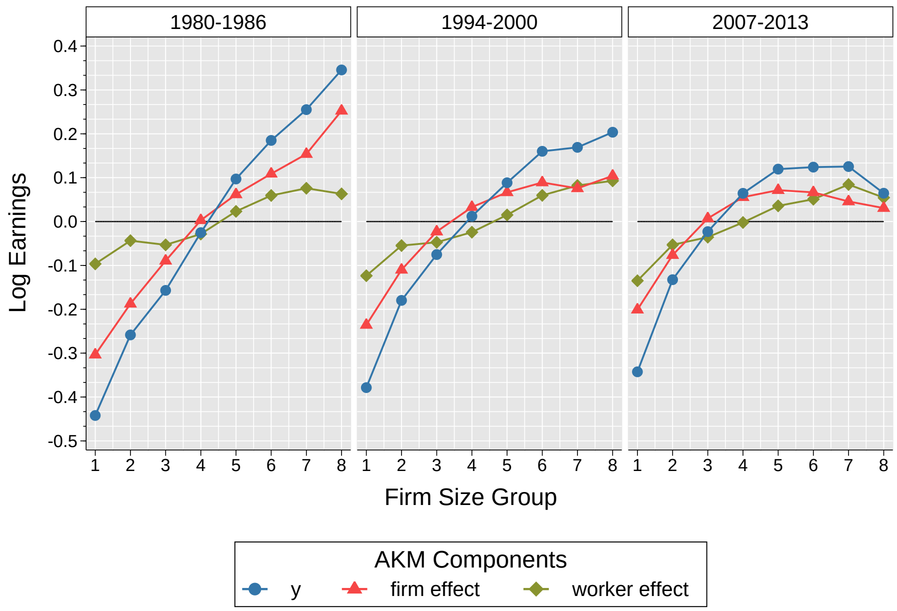
We estimate equation (3.1) separately for five adjacent seven-year intervals beginning in 1980 and ending in 2013. Firm fixed effects are identified by workers moving between firms and hence can only be estimated relative to an omitted firm. Estimation of equation (3.1) is done on the largest set of firms connected by worker flows. To maximize the number of observations in the connected set, we do not impose a restriction on firm size and do not exclude the public sector.[3] Because of limitations in computing power, we estimate equation (3.1) for men only. As we lack data on hours or days of work, our estimates of worker and firm effects may capture systematic differences in labor supply between workers and firms. However, Song et al. (2018) show that a variance decomposition of wage components is robust across a range of labor supply sample restrictions.

3.4 Results

Figure (3.2) presents a visual representation of our main results across three different seven-year intervals. Each panel displays average log earnings in each firm size class (blue line, circles) relative to total average log earnings over the interval. Firms are assigned to eight firm size classes with the smallest firms employing at most ten workers and the largest employing over 15,000. The figure shows that the decline in the large-firm wage premium (LFWP) has not been monotone. On the one hand, the LFWP has been declining strongly at very large employers, starting at 2,500 employees. While in the 1980s going from a 1,000–2,500-employee firm (firm size group 5 in the figure) to a 10,000–15,000-employee firm (group 7) yielded about 15 log points, that gain had completely eroded by 2007–2013. At the bottom of the firm size wage distribution, on the other

[3] Although included in the estimation, public sector jobs are excluded from the empirical analysis.

Figure 3.2: Mean AKM Components by Firm Size and Time Period



Notes: Firm size groups: 1=1-10, 2=10-50, 3=50-250, 4=250-1K, 5=1-2.5K, 6=2.5-10K, 7=10-15K, 8=15K+. Age/year effects and the residual term are omitted.

hand, there is still a sizable earnings premium going from a small firm (say, 10-50, group 2) to a mid-size firm (say, 1,000-2,500). That premium was approximately 35 log points in the 1980s and approximately 25 log points in the last period, with the slight decline arising from an upward shift at the bottom of the firm-size distribution. Since approximately 75% of workers in the U.S. labor market work at firms that are smaller than 2,500 employees, there is still a sizable earnings premium remaining for much of the U.S. workforce.

We also plot the average values of worker and firm earnings components estimated using the AKM estimation equation (3.1) – in particular the firm fixed effect (red line, triangles) and worker fixed effect (olive line, diamonds). Time-variant worker characteristics and the residual component are omitted to highlight the key forces driving the changes over time.

In terms of AKM wage components, the figure shows two key results. First, the major driver of the LFWP in earlier time periods is the firm fixed effect, which accounted for around 70% of the LFWP from 1980 to 1986. That is, the same workers appear to get paid more to work in larger firms. Another 20% of the LFWP is driven by selection effects–workers in larger firms have superior worker fixed effects. The second main finding is that the reduction in the LFWP has almost entirely been driven by the drop in the firm fixed effect premium by firm size. In particular, average earnings have fallen notably for the largest firm size group (15,000+ employees), driven almost entirely by the drop in the firm fixed effect. So, the fall in the LFWP appears to be driven by firms of 1,000 employees or more no longer paying above market salaries to their workers.

Table 3.1: Change in LFWP Regression Coefficients by AKM Components

	Dependent Variable:				
	Log Earnings (1)	Worker Effect (2)	Firm Effect (3)	Age Effect (4)	AKM Residual (5)
Interval 1: 1980-86	0.080	0.016	0.057	0.007	0.001
Interval 5: 2007-13	0.039	0.019	0.021	-0.002	0.001
Change	-0.041	0.003	-0.036	-0.008	0.000
Share (Percent)	-	(-7.5)	(86.8)	(20.2)	(0.5)

In Table (3.1) we formally decompose the change in the LFWP into its constituent AKM wage components. Given equation (3.1), log earnings is additively separable into the AKM components. Therefore, the coefficients in regressions of AKM components on log firm size mechanically add up to the total coefficient of log earnings on log firm size. The decomposition confirms the message of Figure (3.2). The decline in the relationship between firm fixed effects and firm size accounts for 87% of the total decline in the large firm premium. Another factor is a fall in the return to time-varying worker characteristics at large firms–contributing 20% to the total decline in the LFWP. As these characteristics include year and age effects, this result suggests that larger firms are becoming relatively younger. In contrast, selection of worker types by firm size has remained relatively stable over the period. In fact, large firms are slightly more likely to hire high-wage workers in the most recent period. This modest compositional upgrading mitigates the decline of the LFWP–accounting for an 8% increase.

In order to further understand the decline of the LFWP, we turn to an industry analysis. Table (3.2) presents the initial level and changes of both employment and the LFWP by nine broad industries. A few patterns are evident. First, we find a general decline in the LFWP within most industries. In fact, manufacturing is the only industry for which the LFWP did not decline. Second, we find large shifts in employment away from manufacturing, an industry with a high LFWP, into the services sector, an industry with a low LFWP. Industry codes are not assigned to new firms in the SSA data set past the year 2002, therefore, there is also a surge in employment to "unclassified" industries.

Given both within-industry changes in LFWPs and large sectoral shifts in employment, we produce a decomposition to quantify the relative contributions of between- and within-industry factors on the decline in the LFWP. Our main result is that within-industry changes in the LFWP can account for 80% of the decline whereas between-industry factors account for only 20%. (Details of our analysis are in the online appendix). Therefore, the declining LFWP is not merely a reflection of sectoral employment shifts, but suggests broad changes in the pay policies of large firms throughout the economy.

Table 3.2: Change in LFWP and Employment by Industry

	Large Firm Premium 1980-86 (1)	Change in Large Firm Premium (2)	Employment 1980-1986 (millions) (3)	Change in Employment (millions) (4)
Manufacturing	0.094	0.003	85.9	-37.1
Mining	0.104	-0.004	5.6	-2.9
Transportation	0.096	-0.046	26.4	-1.9
Construction	0.095	-0.015	26.0	-1.8
Agriculture	0.049	-0.014	7.1	-1.0
Wholesale Trade	0.060	-0.008	19.7	-0.3
Retail Trade	0.044	-0.051	34.1	2.7
Finance & Insurance	0.057	-0.024	16.5	4.6
Services	0.054	-0.044	53.4	55.3
Unclassified	0.110	-0.048	11.2	79.8

3.5 Conclusion

Large firms have paid a significantly higher wage for more than a century, but over the last thirty years this large-firm wage premium has started to decline. We first document that this reduction is due to a collapse of the wage gradient at very large firms, while the firm size gradient has remained stable for firms with less than 1000–2500 employees. We then show that the decline is largely due a reduction in wage premiums (firm fixed effects) at very large firms, holding worker composition constant. Furthermore the decline cannot be explained by sectoral changes in employment as the majority of the change occurs within industries.

3.6 Appendix: Decomposition of the Change in the LFWP by Industry

We are interested in estimating the large-firm wage premium over time which is the coefficient β_t in the following simple regression model:

$$y_{it} = \alpha_{it} + \beta_t x_{it} + \epsilon_{it}.$$

For random variables x_{it}, y_{it} and industry I_{it} the law of total variance states:

$$Cov\left(x_{it}, y_{it}\right) = E\left[Cov\left(x_{it}, y_{it}|I_{it}\right)\right] + Cov\left(E\left[x_{it}|I_{it}\right], E\left[y_{it}|I_{it}\right]\right).$$

This is the standard between-/within-group variance decomposition where the first term represents the within component and the second term represents the between component. Therefore, we can write the regression coefficient as:

$$\beta_t = \frac{Cov\left(x_{it}, y_{it}\right)}{Var\left(x_{it}\right)} = \frac{E\left[Cov\left(x_{it}, y_{it}|I_{it}\right)\right] + Cov\left(E\left[x_{it}|I_{it}\right], E\left[y_{it}|I_{it}\right]\right)}{E\left[Var\left(x_{it}|I_{it}\right)\right] + Var\left(E\left[x_{it}|I_{it}\right]\right)}.$$

Given that the expression is not additively separable, we propose to decompose the change in the regression coefficient by varying each set of components sequentially. Thus, when assessing the change in β_t between intervals one and five, we create counterfactual regression coefficients for the 5^{th} interval by holding either the between- or within-industry components constant. The order in which the components are varied matters and thus we have a pair of estimates. *Sequence 1* refers to the case with between components change first and then the within components change. *Sequence 2* refers to the opposite case, in which the within components change first. The two sequences provide bounds for the within/between components.

Table (3.3) shows that the large firm premium fell by 4.1 log points between the 1980-1986 interval and the 2007-2013 interval. To put this number into context a worker moving from a 100 employee firm to a 10,000 employee firm would earn 18.9 log points less in 2007-2013 than had he moved in 1980-1986. Panels A and B of Column (1) show that 78 to 80% of the change in the regression coefficient comes through the between-industry components. This result is robust to excluding the unclassified industry. In this case the bounds for the within-industry components range from 73 to 77%. The results are also fairly consistent across intervals with a contribution of the within-industry component of 114%, 90%, and 83% for differences between the 1st and the 2nd, 3rd, and 4th intervals, respectively.

In addition to the decomposition of the total large-firm wage premium, Table (3.3) also provides a decomposition of each of the constituent AKM components of the large-firm wage premium. Note that these components are additively separable as:

$$\begin{aligned}
\beta^y &= \frac{Cov\left(x_{it}, y_{it}\right)}{Var\left(x_{it}\right)} = \frac{Cov\left(x_{it}, \alpha_i + \psi_{j(it)} + x_{it}'\beta + r_{it}\right)}{Var\left(x_{it}\right)} \\
&= \frac{Cov\left(x_{it}, \alpha_i\right)}{Var\left(x_{it}\right)} + \frac{Cov\left(x_{it}, \psi_{j(it)}\right)}{Var\left(x_{it}\right)} + \frac{Cov\left(x_{it}, x_{it}'\beta\right)}{Var\left(x_{it}\right)} + \frac{Cov\left(x_{it}, r_{it}\right)}{Var\left(x_{it}\right)} \\
&= \beta^{\alpha} + \beta^{\psi} + \beta^{x\beta} + \beta^{r}.
\end{aligned}$$

The majority of the change in the large-firm wage premium is due to a reduction in the covariance between firm fixed effects and firm size. In fact, column (3) shows that 87% of the fall in the large firm premium can be attributed to firm fixed effects. Furthermore, in both sequences, changes in firm fixed effects are the key driver of both within- and between-industry reductions in the large-firm wage premium. A secondary factor is a contribution of 20% from the age and year effect components. This is the result of large firms employing a relatively younger workforce. Column (2) shows that worker composition actually works in the opposite direction–responsible for a small rise in the large firm premium. Therefore, although the large premium premium is falling,

Table 3.3: Decomposition of LFWP into Within-/Between-Industry Components

	Dependent Variable:				
	Log Earnings (1)	Worker Effect (2)	Firm Effect (3)	Age Effect (4)	AKM Residual (5)
Panel A: Total Change in Firm Size Regression Coefficient					
Total Change	-0.041	0.003	-0.036	-0.008	0.000
Share (Percent)	-	(-7.5)	(86.8)	(20.2)	(0.5)
Panel B: BT-/WI-Industry Component, Sequence 1					
Within-Industry Change	-0.008	0.005	-0.012	-0.002	0.000
Share (Percent)	(20.0)	(-12.9)	(28.1)	(4.9)	(-0.1)
Between-Industry Change	-0.033	-0.002	-0.024	-0.006	0.000
Share (Percent)	(80.0)	(5.5)	(58.7)	(15.3)	(0.5)
Panel C: BT-/WI-Industry Component, Sequence 2					
Within-Industry Change	-0.009	0.005	-0.012	-0.002	0.000
Share (Percent)	(21.8)	(-12.9)	(29.5)	(5.3)	(-0.1)
Between-Industry Change	-0.032	-0.002	-0.023	-0.006	0.000
Share (Percent)	(78.2)	(5.5)	(57.3)	(15.0)	(0.5)

mean worker quality has slightly improved in large firms. Column (5) shows that the contribution of the residual is negligible.

Bibliography

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

Abowd, John M, Kevin L McKinney, and Ian M Schmutte, "Modeling Endogenous Mobility in Earnings Determination," *Journal of Business & Economic Statistics*, 2017, (Forthcoming).

Abowd, John M., Robert H. Creecy, and Francis Kramarz, "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," Technical Papers 2002-06, Center for Economic Studies, U.S. Census Bureau 2002.

Acemoglu, Daron and David Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," *Handbook of Labor Economics*, 2011, *4*, 1043–1171.

Andrews, Martyn J, Len Gill, Thorsten Schank, and Richard Upward, "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2008, *171* (3), 673–697.

Arrow, Kenneth J., "The Economic Implications of Learning by Doing," *The Review of Economic Studies*, 1962, *29* (3), 155–173.

Atakan, Alp E, "Assortative Matching with Explicit Search Costs," *Econometrica*, 2006, *74* (3), 667–680.

Autor, David, David Dorn, and Gordon H Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *The American Economic Review*, 2013, *103* (6), 2121–2168.

Autor, David H, David Dorn, Gordon H Hanson, and Jae Song, "Trade Adjustment: Worker-Level Evidence," *The Quarterly Journal of Economics*, 2014, *129* (4), 1799–1860.

_, Lawrence F Katz, and Melissa S Kearney, "Trends in US Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, 2008, *90* (2), 300–323.

Bagger, Jesper and Rasmus Lentz, "An Empirical Model of Wage Dispersion with Sorting," *The Review of Economic Studies*, 2018.

_, Kenneth L Sørensen, and Rune Vejlin, "Wage Sorting Trends," *Economics Letters*, 2013, *118* (1), 63–67.

Baker, George, Michael Gibbs, and Bengt Holmstrom, "The Internal Economics of the Firm: Evidence from Personnel Data," *The Quarterly Journal of Economics*, 1994, pp. 881–919.

_, _, and _, "The Wage Policy of a Firm," *The Quarterly Journal of Economics*, 1994, pp. 921–955.

Bartolucci, Cristian, Francesco Devicienti, and Ignacio Monzón, "Identifying Sorting in Practice," Technical Report, Institute for the Study of Labor (IZA) 2015.

Baumgarten, Daniel, "Exporters and the Rise in Wage Inequality: Evidence from German Linked Employer–Employee Data," *Journal of International Economics*, 2013, *90* (1), 201–217.

Beaudry, Paul and John DiNardo, "The Effect of Implicit Contracts on the Movement of Wages Over the Business Cycle: Evidence from Micro Data," *Journal of Political Economy*, 1991, pp. 665–688.

Becker, Gary S, *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, The National Bureau of Eocnomic Research, 1964.

_ , "A Theory of Marriage: Part I," *Journal of Political Economy*, 1973, *81* (4), 813–846.

Ben-Porath, Yoram, "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 1967, *75* (4), 352–365.

Bertrand, Marianne and Sendhil Mullainathan, "Are CEOs Rewarded for Luck? The Ones Without Principals Are," *Quarterly Journal of Economics*, 2001, pp. 901–932.

Blinder, Alan S, "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources*, 1973, pp. 436–455.

Bloom, Nick, Tobias Kretschmer, and John Van Reenen, "Are Family-Friendly Workplace Practices a Valuable Firm Resource?," *Strategic Management Journal*, 2011, *32* (4), 343–367.

Bombardini, Matilde, Gianluca Orefice, and Maria D Tito, "Does Exporting Improve Matching? Evidence from French Employer-Employee Data," Technical Report, National Bureau of Economic Research 2015.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "A Distributional Framework for Matched Employer Employee Data," *Unpublished manuscript, University of Chicago*, 2016.

Borovicková, Katarına and Robert Shimer, "High Wage Workers Work for High Wage Firms," *Unpublished Manuscript*, 2017.

Brown, Charles and James Medoff, "The Employer Size-Wage Effect," *Journal of Political Economy*, 1989, *97* (5), 1027–1059.

Brown, Jennifer, "Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars," *Journal of Political Economy*, 2011, *119* (5), 982–1013.

Burdett, Kenneth and Dale T. Mortensen, "Wage Differentials, Employer size, and Unemployment," *International Economic Review*, 1998, pp. 257–273.

Bustos, Paula, "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms," *The American Economic review*, 2011, *101* (1), 304–340.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

_ , _ , and _ , "Optimal Data-Driven Regression Discontinuity Plots," *Journal of the American Statistical Association*, 2015, *110* (512), 1753–1769.

_ , _ , Max H Farrell, and Rocıo Titiunik, "Regression Discontinuity Designs Using Covariates," *Unpublished, University of Michigan*, 2016.

_ , _ , _ , and _ , "rdrobust: Software for Regression Discontinuity Designs," *Stata Journal, forthcoming*, 2017.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women," *The Quarterly Journal of Economics*, 2015, *131* (2), 633–686.

_ , _ , Jörg Heining, and Patrick Kline, "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

_ and Thomas Lemieux, "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *The Quarterly Journal of Economics*, 2001, *116* (2), 705–746.

_ , Jörg Heining, and Patrick Kline, "Workplace Heterogeneity and the Rise of West German Wage Inequality," *The Quarterly Journal of Economics*, 2013, *128* (3), 967–1015.

Carrington, William J and Bruce Fallick, “Why Do Earnings Fall with Job Displacement?,” *Industrial Relations: A Journal of Economy and Society*, 2017, *56* (4), 688–722.

__ , Kristin McCue, and Brooks Pierce, “The Role of Employer/Employee Interactions in Labor Market Cycles: Evidence from the Self-Employed,” *Journal of labor Economics*, 1996, *14* (4), 571–602.

Cattaneo, Matias D, Brigham R Frandsen, and Rocio Titiunik, “Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the US Senate,” *Journal of Causal Inference*, 2015, *3* (1), 1–24.

__ , Michael Jansson, and Xinwei Ma, “rddensity: Manipulation Testing Based on Density Discontinuity,” *The Stata Journal*, 2016, pp. 1–18.

__ , __ , and __ , “Simple Local Regression Distribution Estimators with an Application to Manipulation Testing,” Technical Report, University of Michigan Working Paper 2016.

__ , Rocio Titiunik, and Gonzalo Vazquez-Bare, “Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality,” Technical Report, Working Paper, University of Michigan 2016.

__ , __ , and __ , “rdlocrand: Inference in Regression Discontinuity Designs under Local Randomization,” *Stata Journal, forthcoming*, 2017.

Couch, Kenneth A and Dana W Placzek, “Earnings Losses of Displaced Workers Revisited,” *The American Economic Review*, 2010, *100* (1), 572–589.

Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum, “The Rise of the East and the Far East: German Labor Markets and Trade Integration,” *Journal of the European Economic Association*, 2014, *12* (6), 1643–1675.

___ , ___ , and ___ , "Adjusting to Globalization-Evidence from Worker-Establishment Matches in Germany," Technical Report, CEPR Discussion Paper No. DP11045 2016.

Davidson, Carl, Fredrik Heyman, Steven Matusz, Fredrik Sjöholm, and Susan Chun Zhu, "Globalization and Imperfect Labor Market Sorting," *Journal of International Economics*, 2014, *94* (2), 177–194.

___ , Steven J Matusz, and Andrei Shevchenko, "Globalization and Firm Level Adjustment with Imperfect Labor Markets," *Journal of International Economics*, 2008, *75* (2), 295–309.

Davis, Steven J and Till von Wachter, "Recessions and the Costs of Job Loss," *Brookings Papers on Economic Activity*, 2011, *43* (2 (Fall)), 1–72.

de Melo, Rafael Lopes, "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," *Journal of Political Economy*, 2018, *126* (1), 000–000.

Deaton, Angus, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, World Bank Publications, 1997.

DeBacker, Jason, Bradley Heim, Vasia Panousi, Shanthi Ramnath, and Ivan Vidangos, "Rising Inequality: Transitory or Persistent? New Evidence from a Panel of US Tax Returns," *Brookings Papers on Economic Activity*, 2013, *2013* (1), 67–142.

Delacroix, Alain and Shouyong Shi, "Directed Search on the Job and the Wage Ladder," *International Economic Review*, 2006, *47* (2), 651–699.

DiNardo, John, Nicole M Fortin, and Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," Technical Report, National Bureau of Economic Research 1995.

Dustmann, Christian and Albrecht Glitz, "How Do Industries and Firms Respond to Changes in Local Labor Supply?," *Journal of Labor Economics*, 2015, *33* (3 Part 1), 711–750.

_ **, Johannes Ludsteck, and Uta Schönberg**, "Revisiting the German Wage Structure," *The Quarterly Journal of Economics*, 2009, *124* (2), 843–881.

Eeckhout, Jan and Philipp Kircher, "Identifying Sorting–In Theory," *The Review of Economic Studies*, 2011.

Fallick, Bruce and Charles A Fleischman, "Employer-to-Employer Flows in the US labor market: The Complete Picture of Gross Worker Flows," Technical Report, FEDS Working Paper No. 2004-34 2004.

Fallick, Bruce C, "A Review of the Recent Empirical Literature on Displaced Workers," *ILR Review*, 1996, *50* (1), 5–16.

Farber, Henry S, "The Incidence and Costs of Job Loss: 1982-91," *Brookings Papers on Economic Activity. Microeconomics*, 1993, *1993* (1), 73–132.

_ , "The Changing Face of Job Loss in the United States, 1981-1993," Technical Report, National Bureau of Economic Research 1996.

_ , "Job Loss in the Great Recession and Its Aftermath: US Evidence from the Displaced Workers Survey," Technical Report, National Bureau of Economic Research 2015.

_ , "Employment, Hours, and Earnings Consequences of Job Loss: US Evidence from the Displaced Workers Survey," *Journal of Labor Economics*, 2017, *35* (S1), S235–S272.

Feinstein, John, *Tales from Q School: Inside Golf's Fifth Major*, Little, Brown, 2007.

Genda, Yuji, Ayako Kondo, and Souichi Ohta, "Long-Term Effects of a Recession at Labor Market Entry in Japan and the United States," *Journal of Human Resources*, 2010, *45* (1), 157–196.

Gibbons, Robert and Lawrence F Katz, "Layoffs and Lemons," *Journal of Labor Economics*, 1991, pp. 351–380.

_ **and Michael Waldman**, "Enriching a Theory of Wage and Promotion Dynamics inside Firms," *Journal of Labor Economics*, 2006, *24* (1), 59–108.

Goldschmidt, Deborah and Johannes F Schmieder, "The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure," *The Quarterly Journal of Economics*, 2017, p. qjx008.

Gottschalk, Peter and Robert Moffitt, "The Growth of Earnings Instability in the US Labor Market," *Brookings Papers on Economic Activity*, 1994, *1994* (2), 217–272.

Goux, Dominique and Eric Maurin, "Persistence of Interindustry Wage Differentials: A Reexamination Using Matched Worker-Firm Panel Data," *Journal of Labor Economics*, 1999, *17* (3), 492–533.

Greenwald, Bruce C, "Adverse Selection in the Labour Market," *The Review of Economic Studies*, 1986, *53* (3), 325–347.

Gruber, Jonathan and James Poterba, "Tax Incentives and the Decision to Purchase Health Insurance: Evidence from the Self-Employed," *The Quarterly Journal of Economics*, 1994, *109* (3), 701–733.

Gruetter, Max and Rafael Lalive, "The Importance of Firms in Wage Determination," *Labour Economics*, 2009, *16* (2), 149–160.

Gulyas, Andreas, "Identifying Labor Market Sorting with Firm Dynamics," Technical Report, Mimeo, UCLA 2016.

Guryan, Jonathan, Kory Kroft, and Matthew J Notowidigdo, "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments," *American Economic Journal: Applied Economics*, 2009, *1* (4), 34–68.

Guvenen, Fatih, Greg Kaplan, Jae Song, and Justin Weidner, "Lifetime Incomes in the United States over Six Decades," Technical Report, National Bureau of Economic Research 2017.

Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii, "Identifying Equilibrium Models of Labor Market Sorting," *Econometrica*, 2017, *85* (1), 29–65.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 2001, *69* (1), 201–209.

Håkanson, Christina, Erik Lindqvist, and Jonas Vlachos, "Firms and Skills: The Evolution of Worker Sorting," Technical Report, Working Paper, IFAU-Institute for Evaluation of Labour Market and Education Policy 2015.

Hall, Robert E, "Technical Change and Capital from the Point of View of the Dual," *The Review of Economic Studies*, 1968, *35* (1), 35–46.

Haltiwanger, John, Henry Hyatt, and Erika McEntarfer, "Do Workers Move Up the Firm Productivity Job Ladder?," *Unpublished Manuscript*, 2017.

_ , _ , and _ , "Who Moves Up the Job Ladder?," Technical Report, National Bureau of Economic Research 2017.

Hamermesh, Daniel S, "What Do We Know About Worker Displacement in the US?," *Industrial Relations: A Journal of Economy and Society*, 1989, *28* (1), 51–59.

Heckman, James J, Lance Lochner, and Christopher Taber, "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1998, *1* (1), 1–58.

Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding, "Inequality and Unemployment in a Global Economy," *Econometrica*, 2010, *78* (4), 1239–1283.

Hethey, Tanja and Johannes F Schmieder, "Using Worker Flows in the Analysis of Establishment Turnover–Evidence from German Ddministrative Data," *FDZ Methodenreport*, 2010, *6* (en), 43.

Hipple, Steven F, "Self-Employment in the United States," *Monthly Labor Review*, 2010, *133* (9), 17–32.

Holzer, Harry J, Julia I Lane, David B Rosenblum, and Fredrik Andersson, *Where Are All the Good Jobs Going?: What National and Local Job Quality and Dynamics Mean for U.S. Workers*, Russell Sage Foundation, 2011.

Hornstein, Andreas, Per Krusell, and Giovanni L Violante, "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *The American Economic Review*, 2011, *101* (7), 2873–98.

Huggett, Mark, Gustavo Ventura, and Amir Yaron, "Sources of Lifetime Inequality," *The American Economic Review*, 2011, *101* (7), 2923–54.

Iranzo, Susana, Fabiano Schivardi, and Elisa Tosetti, "Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data," *Journal of Labor Economics*, 2008, *26* (2), 247–285.

Jarosch, Gregor, "Searching for Job Security and the Consequences of Job Loss," *Manuscript, Princeton University*, 2015.

Jung, Philip and Moritz Kuhn, "Earnings Losses and Labor Mobility Over the Life Cycle," *Journal of the European Economic Association*, 2018.

Kahn, Lisa B, "The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy," *Labour Economics*, 2010, *17* (2), 303–316.

Kambourov, Gueorgui and Iourii Manovskii, "Occupational Specificity of Human Capital," *International Economic Review*, 2009, *50* (1), 63–115.

Katz, Lawrence F and David H Autor, "Changes in the Wage Structure and Earnings Inequality," *Handbook of Labor Economics*, 1999, *3*, 1463–1555.

_ **and Lawrence H Summers**, "Industry rents: Evidence and implications," *Brookings Papers on Economic Activity. Microeconomics*, 1989, *1989*, 209–290.

_ **, Gary W Loveman, and David G Blanchflower**, "A Comparison of Changes in the Structure of Wages," Technical Report, National Bureau of Economic Research 1993.

Kletzer, Lori G, "Job Displacement," *Journal of Economic perspectives*, 1998, *12* (1), 115–136.

Krashinsky, Harry, "The Effect of Labor Market Institutions on Salaried and Self-Employed Less-Educated Men in the 1980s," *ILR Review*, 2008, *62* (1), 73–91.

Krolikowski, Pawel, "Job Ladders and Earnings of Displaced Workers," *American Economic Journal: Macroeconomics*, 2017, *9* (2), 1–31.

Krueger, Alan B and Lawrence H Summers, "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica: Journal of the Econometric Society*, 1988, pp. 259–293.

Lachowska, Marta, Alexandre Mas, and Stephen A Woodbury, "Sources of Displaced Workers Long-Term Earnings Losses," Technical Report, National Bureau of Economic Research 2018.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman, "Life-Cycle Wage Growth Across Countries. Forthcoming," *Journal of Political Economy*, 2016.

Lazear, Edward P, "Agency, Earnings Profiles, Productivity, and Hours Restrictions," *The American Economic Review*, 1981, *71* (4), 606–620.

__ **and Robert L Moore**, "Incentives, Productivity, and Labor Contracts," *The Quarterly Journal of Economics*, 1984, *99* (2), 275–296.

Lee, David S and David Card, "Regression Discontinuity Inference with Specification Error," *Journal of Econometrics*, 2008, *142* (2), 655–674.

Lileeva, Alla and Daniel Trefler, "Improved Access to Foreign Markets Raises Plant-Level Productivity For Some Plants," *The Quarterly Journal of Economics*, 2010, *125* (3), 1051–1099.

Lise, Jeremy and Jean-Marc Robin, "The Macrodynamics of Sorting between Workers and Firms," *The American Economic Review*, 2017, *107* (4), 1104–1135.

__ **, Costas Meghir, and Jean-Marc Robin**, "Matching, Sorting and Wages," *Review of Economic Dynamics*, 2016, *19*, 63–87.

MacLeod, W Bentley and James M Malcomson, "Investments, Holdup, and the Form of Market Contracts," *The American Economic Review*, 1993, pp. 811–837.

Maestas, Nicole, Kathleen J Mullen, David Powell, Till von Wachter, and Jeffrey B Wenger, "Working Conditions in the United States," 2017.

McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

Melitz, Marc J, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

Mincer, Jacob, *Schooling, Experience, and Earnings*, New York: Columbia University Press, 1974.

Moffitt, Robert A and Peter Gottschalk, "Trends in the Transitory Variance of Male Earnings in the US, 1970-2004," Technical Report, National Bureau of Economic Research 2011.

Moore, Henry L., *Laws of wages: An essay in statistical economics*, New York: The Macmillan Company, 1911.

Moore, Robert L, "Employer Discrimination: Evidence from Self-Employed Workers," *The Review of Economics and Statistics*, 1983, pp. 496–501.

Müller, Steffen, Jens Stegmaier, and Moises Yi, "Industry Mix, Local Labor Markets, and the Incidence of Trade Shocks," 2016.

Nagypál, Éva, "Worker Reallocation over the Business Cycle: The Importance of Employer-to-Employer Transitions," *Manuscript, Northwestern University*, 2008.

Neal, Derek, "Industry-Specific Human Capital: Evidence from Eisplaced Workers," *Journal of Labor Economics*, 1995, pp. 653–677.

Oaxaca, Ronald, "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 1973, pp. 693–709.

Oi, Walter Y. and Todd L. Idson, "Firm Size and Wages," *Handbook of Labor Economics*, 1999, *3, Part B*, 2165 – 2214.

Olsen, Anya and Russell Hudson, "Social Security Administration's Master Earnings File: Background Information," *Social Security Bulletin*, 2009, *69* (3), 29.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz, "The Short- and Long-Term Career Effects of Graduating in a Recession," *American Economic Journal: Applied Economics*, January 2012, *4* (1), 1–29.

Oyer, Paul, "Initial Labor Market Conditions and Long-Term outcomes for Economists," *The Journal of Economic Perspectives*, 2006, *20* (3), 143–160.

_ , "The Making of an Investment Banker: Stock Market Shocks, Career Choice, and Lifetime Income," *The Journal of Finance*, 2008, *63* (6), 2601–2628.

Parker, Simon C, "The Distribution of Self-Employment Income in the United Kingdom, 1976–1991," *The Economic Journal*, 1997, *107* (441), 455–466.

_ , *The Economics of Self-Employment and Entrepreneurship*, Cambridge University Press, 2004.

_ , Yacine Belghitar, and Tim Barmby, "Wage Uncertainty and the Labour Supply of Self-Employed Workers," *The Economic Journal*, 2005, *115* (502).

Pope, Devin G and Maurice E Schweitzer, "Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes," *The American Economic Review*, 2011, *101* (1), 129–157.

Rosen, Sherwin, "Learning and Experience in the Labor Market," *Journal of Human Resources*, 1972, pp. 326–342.

_ , "Learning by Experience as Joint Production," *The Quarterly Journal of Economics*, 1972, pp. 366–382.

Ruhm, Christopher J, "Are Workers Permanently Scarred by Job Displacements?," *The American Economic Review*, 1991, *81* (1), 319–324.

Salop, Joanne and Steven Salop, "Self-Selection and Turnover in the Labor Market," *The Quarterly Journal of Economics*, 1976, pp. 619–627.

Sampson, Thomas, "Selection into Trade and Wage Inequality," *American Economic Journal: Microeconomics*, 2014, *6* (3), 157–202.

Schank, Thorsten, Claus Schnabel, and Joachim Wagner, "Do Exporters Really Pay Higher Wages? First Evidence from German Linked Employer–Employee Data," *Journal of International Economics*, 2007, *72* (1), 52–74.

Schmutte, Ian M, "Job Referral Networks and the Determination of Earnings in Local Labor Markets," *Journal of Labor Economics*, 2014, *33* (1), 1–32.

Shimer, Robert and Lones Smith, "Assortative Matching and Search," *Econometrica*, 2000, *68* (2), 343–369.

Smith, Benjamin S., Jae Song, and Till von Wachter, "Earnings Growth, Job Mobility, and Inequality from 1957 to 2014," Technical Report, Unpublished Manuscript 2017.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter, "Firming Up Inequality," Technical Report, Federal Reserve Bank of Minneapolis 2018.

Stevens, Ann Huff, "Persistent Effects of Job Displacement: The Importance of Multiple Job Losses," *Journal of Labor Economics*, 1997, *15* (1, Part 1), 165–188.

Sullivan, Daniel and Till von Wachter, "Job Displacement and Mortality: An Analysis Using Administrative Data," *The Quarterly Journal of Economics*, 2009, *124* (3), 1265–1306.

Topel, Robert, "Specific Capital and Unemployment: Measuring the Costs and Consequences of Job Loss," in "Carnegie-Rochester Conference Series on Public Policy," Vol. 33 Elsevier 1990, pp. 181–214.

Topel, Robert H and Michael P Ward, "Job Mobility and the Careers of Young Men," *The Quarterly Journal of Economics*, 1992, pp. 439–479.

von Wachter, Till, "Summary of the Literature on Job Displacement in the US and EU: What We Know and What We Would Like to Know," *David Marsden and François Ryxc (eds.), Wage Structures, Employment Adjustments and Globalization: Evidence from Linked and Firm-level Panel Data, Applied Econometrics Association Series, Palgrave Macmillan*, 2009.

_ , "Measuring the Short- and Long-Term Cost of Job Displacement in Terms of Worker and Family Outcomes," Technical Report, Working Paper, UCLA 2015.

_ **, Jae Song, and Joyce Manchester**, "Long-Term Earnings Losses Due to Mass Layoffs During the 1982 Recession: An Analysis Using US Administrative Data from 1974 to 2004," Working Paper, Institute for the Study of Labor September 2009.

Woodcock, Simon D, "Match Effects," *Research in Economics*, 2015, *69* (1), 100–121.