Essays on Labor Economics and Advertising Auctions

by

Joseph M. Golden

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2014

Doctoral Committee:

Professor John Bound, Chair
Professor Charles C. Brown
Assistant Professor Eric M. Schwartz
Professor Frank P. Stafford

ACKNOWLEDGMENTS

I am grateful to the many people who helped me complete my dissertation. First, I think my family. I could not have made it through graduate school without Lindsey's kindness, generosity and willingness to proofread my work. My parents, Sandy and Ted, both have been tremendously supportive of my education and work. They encouraged me to learn and to pursue my passions. My siblings, Amy, Steve, David, Allan and Melissa are always there for me and have been wonderful to talk to about my research.

My advisers have encouraged my work and I am thankful for their support. John Bound, the chair of my dissertation committee, taught me a enormous amount as an adviser, mentor, co-author and friend. Charlie Brown provided excellent feedback on all of my papers, especially on the second chapter of my dissertation. I took Frank Stafford's class as an undergraduate and he encouraged me to pursue graduate school. His feedback and support of my research have been invaluable. Eric Schwartz provided excellent guidance and feedback on the first chapter of my dissertation.

Friends, co-authors and additional advisers have all supported me at various stages of graduate school. When I was an undergraduate, Gary Solon encouraged me to become an economist. I thank Lones Smith for sharing his love of theory with me. Sarah Turner provided helpful career guidance and mentorship. In addition to John Bound, Breno Graga and Gaurav Khanna were excellent co-authors on the third chapter of my dissertation. Dan Hirschman debated ideas with me as a housemate and longtime friend. Kevin Borders encouraged me to pursue a Ph.D. and I have learned many things from him. Evan Herrnstadt, Jason Kerwin and Ophira Vishkin provided frequent feedback and support during our regular video conferences which started when we were all away from our department in various locations. Isaac Sorkin provided feedback on my research and was a co-author on earlier versions of my second chapter. Tom Blake, John Horton and Steve Tadelis have helped me with various stages of the first chapter of my dissertation. I also thank Hal Varian and Qing Wu for everything I learned at Google, and Ved Sinha and Servaes Tholen for supporting my work at Elance-oDesk.

TABLE OF CONTENTS

Acknowledgments . . . . **ii**

List of Figures . . . . **v**

List of Tables . . . . **vi**

Chapter

1 Search Advertising Effects on Competitors:
An Experiment Before a Merger **1**

1.1 Introduction 1
1.2 Legal and Market Context 5
1.2.1 Legal Context of Ads on Trademarked Search Terms 5
1.2.2 Market Context 7
1.3 Experiment Design 7
1.4 Hypotheses Development 9
1.4.1 Effects on oDesk's Search Ad Campaign 9
1.4.2 Effects on Elance and oDesk Business 11
1.5 Empirical Methods 12
1.6 Results 13
1.6.1 The Experiment's Effects on oDesk's Advertising Campaign 14
1.6.2 The Experiment's Effects on Elance's and oDesk's Businesses 15
1.7 Conclusion 17
A1 Description of Search Ad Auctions 30

2 The Importance of Preferences in Sectoral Sorting:
Direct Evidence from Lawyers in the Private and Nonprofit Sectors **33**

2.1 Introduction 33
2.2 Data 37
2.2.1 Sectoral Definition 37
2.2.2 Summary Statistics 38
2.2.3 Key Features of the Data 39
2.3 A Conceptual Model of Sorting 43
2.3.1 Environment 43
2.3.2 Equilibrium: Sorting and Selection 44

2.4 Main Results . . . 45
2.4.1 Empirical Model of Private Sector Wages . . . 45
2.4.2 Wage Equation Estimates By Sector . . . 46
2.4.3 Estimates of Opportunity Cost of Working in the Nonprofit Sector . . . 46
2.4.4 Robustness of Estimated Opportunity Cost to Selection on Unobservables 48
2.4.5 The Role of Preferences Within the Private Sector . . . 50
2.5 Fixed Effect Analysis of Selection . . . 51
2.5.1 Analysis of Sector Switchers . . . 51
2.5.2 Change of Sector Plans During Law School . . . 55
2.6 Alternative Potential Explanations for the Sectoral Pay Gap . . . 56
2.6.1 Lifetime Earnings . . . 56
2.6.2 Debt Forgiveness . . . 57
2.6.3 Differences in Family Background . . . 59
2.7 Conclusion . . . 60
B1 Detailed Data Description . . . 83
B1.1 After the JD . . . 83
B1.2 University of Michigan Law School Survey . . . 86

3 Recruitment of Foreigners in the Market for Computer Scientists in the US . . . 93

3.1 Introduction . . . 93
3.2 The Market for Computer Scientists in the 1990s . . . 96
3.2.1 The Information Technology Boom of the Late 1990s . . . 96
3.2.2 The Immigrant Contribution to the Growth of the High Tech Workforce . 98
3.2.3 The Previous Literature on the Impact of Immigrants on the High Tech Workforce in the US . . . 100
3.3 A Dynamic Model of Supply and Demand of Computer Scientists . . . 103
3.3.1 Labor Supply of American Computer Scientists . . . 104
3.3.2 Labor Supply of Foreign Computer Scientists . . . 108
3.3.3 Labor Demand for Computer Scientists . . . 109
3.3.4 Equilibrium . . . 113
3.4 Calibration and Simulation . . . 114
3.4.1 Identification and Calibration Method . . . 114
3.4.2 Calibration results . . . 116
3.4.3 Simulation of Fixed Foreign Worker Population Counterfactual . . . 118
3.4.4 Identification of Labor Demand . . . 120
3.5 Discussion . . . 122

Bibliography . . . 131

LIST OF FIGURES

1.1 Google search results including oDesk and Elance ads and organic results 26
1.2 oDesk search ad campaign for the term “Elance” 27
1.3 oDesk search ad campaign, overall 28
1.4 Elance and oDesk new client registrations 29

2.1 Distribution of Private Sector Annual Compensation by Law School Rank in the Private and Nonprofit Sectors 78
2.2 Preferences and Sector Choice 79
2.3 Preferences and Skill 80
2.4 Conceptual Diagram of Sorting by Skill and Preferences 81
2.5 Lawyer Pay, GPA and Sectoral Sorting Over Time 82

3.1 Major Trends (1990 to 2012) 127
3.2 H-1 and H-1B Visa Population 128
3.3 Model and Counterfactual (1/2) 129
3.4 Model and Counterfactual (2/2) 130

LIST OF TABLES

1.1 Game models of the search ad auction when a trademark owner bids against a competitor, with best responses in bold. 19
1.2 Description of advertising metrics and predicted effects of the experiment on oDesk's Google ad campaign. 20
1.3 Description of Elance and oDesk business metrics and predicted effects of the experiment on them. 21
1.4 Mean variable values before and during the experiment, by experiment cell, with differences and difference-in-differences 22
1.5 oDesk ads on "Elance" term estimates 23
1.6 oDesk ads entire campaign estimates 24
1.7 New user registrations estimates 25

2.1 Lawyer summary statistics by sector, After the JD 61
2.2 Summary statistics by sector, University of Michigan Law School 62
2.3 Sectoral choice probit marginal effects. 63
2.4 Work preferences by sector (AJD 2002) 64
2.5 Work preferences by law school rank 65
2.6 AJD Private Sector Log Hourly Wage Regressions 66
2.7 AJD Nonprofit Sector Log Hourly Wage Regressions 67
2.8 UMLS log hourly wage regressions by sector 68
2.9 Estimated average opportunity cost of working in the nonprofit sector 69
2.10 Estimated average opportunity cost of working in the nonprofit sector by law school rank 70
2.11 Sensitivity of estimated average opportunity cost of working in the nonprofit sector to unobservables 71
2.12 Sector Switchers from AJD 72
2.13 AJD switchers: log hourly wage regression 73
2.14 Sector Switchers from UMLS 74
2.15 UMLS switchers: log annual wage regressions 75
2.16 Sector Plan Switchers from the UMLS 76
2.17 Debt by Law School Rank 77
B1 Construction of the Sector Variable (AJD 2002) 89
B2 UMLS Survey Years 90
B3 Law School Financial Support Percentages by Sector 91

B4 Parents' Education by Sector, AJD 92

3.1 Fraction of Computer Scientists and Immigrants in the US Workforce by Occupation . 124
3.2 Calibrated Parameters 125
3.3 Summary of Results from Counterfactual Simulation 126

CHAPTER 1

Search Advertising Effects on Competitors: An Experiment Before a Merger

1.1 Introduction

Elance and oDesk are the two largest online work marketplaces. They connect employers and freelancer workers and provide software tools to facilitate them working together online. On April 1, 2014, they merged. Prior to merging, these companies were fierce business competitors, and battling search advertisers, often targeting their ads at the same search terms on Google and other search engines. Figure 1.1 shows an example screenshot of a search engine results page for a search term where both Elance and oDesk ads appear. Before the merger, we conducted an experiment where we shut down Elance's search advertising in half of the United States, randomized at the regional level.[1] Following the conclusion of the experiment and the consummation of the merger, we gained access to oDesk's advertising data and internal databases to assess the impact of the Elance experiment on oDesk. Access to data from two search advertising and business competitors is unique to our setting and central to our analysis.

We examine two sets of experimental effects: (1) the effects on oDesk's search advertising campaign and business, which we assess by exploiting the special opportunity resulting from the merger, with a particular focus on the competition between the two companies over search ads on their trademarked search keywords, and (2) the efficacy of Elance's ads as measured by the

[1] Joseph Golden was an employee of Elance prior to the merger and designed the experiment.

effects on its own business (which we compare to a naive observational estimate). The difficulty of estimating these effects observationally motivated conducting an experiment. The most surprising result is that although during the experiment, on searches for the term "Elance," Elance's ad received 100 times as much traffic as oDesk's, with Elance's ad out of the way, oDesk did not gain substantial additional clicks.

In the United States, online advertising experts predict that in 2014, companies will spend $20 billion on online brand advertising (ads designed to build consumer awareness, interest and loyalty in their brands) and $30 billion on online direct response advertising (ads designed to elicit a specific action such as an online purchase).[2] Advertisers purchase sponsored search ads, the ads that appear along with search results on search engines including Google and Bing, primarily as direct response marketing, but also for branding purposes. Additionally, advertisers defensively purchase search ads to block their competitors' ads from appearing or being featured prominently. One reason to do so is if competitors gaining business now would hurt an advertiser's future prospects, such as if a market has network effects or a winner-take-all dynamics.

Search engines including Google sell their ads via real-time auctions in which advertisers bid on search terms.[3] Companies may generally bid on ads for their own and their competitors' trademarked terms. Search engine marketing experts claim that trademark owners and competitors play a version of the Prisoner's Dilemma (PD) when bidding on trademarked search ads, and that owners must bid aggressively to block competitors from poaching their potential traffic.[4] However, advertisers cannot easily verify this claim and optimize their advertising campaigns accordingly for two reasons. First, they need a sufficient amount of exogenous variation in advertising behavior. Second, they need data from competing businesses. Using our combination of experimental evidence and novel data sets from Elance and oDesk, our results suggest that the trademark

[2] According to eMarketer, a large online market research firm: `http://www.emarketer.com/Article/Direct-Response-Tactics-Take-Majority-of-US-Marketers-Budgets/1010852`. Accessed August 8, 2014.

[3] Appendix A1 describes search ad auctions in more detail.

[4] For example, see this blog post at Search Engine Land, a popular blog and resource for search engine marketing professions, which calls allowing a competitor to outbid you on your own trademarked terms an "obviously untenable situation": `http://searchengineland.com/how-to-protect-brand-keywords-for-less-121566`. Similarly, in the legal literature, Gervais et al. (2013) argues that trademark owners bid on their own terms to block their competition.

owner/competitor game is not in fact a PD.

Table 1.1 shows two possible, simple game structures that we consider as candidates to model the trademark owner (O) and competitor (C) search ad bidding game. The first, in Table 1.1a, demonstrates the PD game in which bidding is a dominant strategy for both players. Its story goes as follows: if competing O and C both bid on search ads for O's trademarked terms, the auctioneer benefits by earning more revenue than if either company instead does not participate. The companies could collude and not bid, each saving on advertising expenses to their mutual benefit, but if O decides to stop bidding on its own terms, then C will poach traffic from O's terms by bidding, making unilateral deviation on O's part a bad move. Similarly, if C decides to not bid, O will get an increased share of the traffic, making unilateral deviation on C's part a bad move as well. The second, in Table 1.1b, demonstrates a different game with a single difference: poaching is less effective, and so O has less to lose by not bidding (and C thus gains less). Hence, O does not bid in this game's Nash Equilibrium. Our experimental evidence suggests that many advertisers are likely playing a game more accurately described by Table 1.1b, where bidding is less effective than they think at blocking their competition, and they should thus not bid on their own trademarked terms.

Prior to merging, Elance and oDesk typically each occupied the top advertising position in Google searches for its own name ("Elance" and "oDesk" respectively), and usually occupied the second position for the other company's name. Both businesses target their search ads at potential new employers. During our experiment, as the experimental treatment, Elance unilaterally stopped bidding on all of its Google search ads, including all of its ads on both general and trademarked terms, in half of the United States, randomized geographically.[5] Any company which is currently bidding on its own trademarked terms could estimate the savings they would attain by stopping their ads. However, to estimate a key presumed cost of not bidding, lost business to competitors, a company would need both exogenous variation in their own advertising on these terms and access to their competitors' data. We meet both of these conditions since we introduced exogenous

[5]Many websites including Google and Elance use IP Geolocation software to estimate the location of their users based on their IP addresses. Elance did not geographically target its other non-Google ads, including its smaller search ad campaigns on other search engines, towards or away from the treated areas in the experiment.

variation via our experiment, which we conducted before Elance and oDesk combined their online advertising operations, and we analyze the results of the experiment after the companies merged, using data from each company's advertising campaign and relevant business metrics.

Prior to the experiment and during the experiment in the regions where Elance continued to purchase Google search ads, Google searches for the term "Elance" would generally show an ad for Elance in the top position and an ad for oDesk in the second position. Users making these searches only clicked on oDesk's ad about 1% as often as they clicked on Elance's ad. This extremely low percentage may indicate that these searches were typically "navigational" in nature, meaning that users searching for "Elance" appear to have generally been using Google to navigate to their intended destination website, as opposed to searching to explore or discover different websites.[6] When the companies were competitors, Elance of course did not know that oDesk received relatively few clicks on its ads targeting the "Elance" term. However, the fact that there were few clicks on oDesk's ads suggests, misleadingly, that Elance had a lot of downside risk if they chose to not bid, because oDesk could have potentially captured a large amount of Elance's traffic from the "Elance" term by outbidding Elance. When Elance turned off its own ads in half of the U.S. during the experiment, oDesk's ads on the term "Elance" moved into the top advertising position. However, surprisingly, oDesk received nearly the same number of clicks on its ads in the half of the U.S. where Elance shut its ads off as in the half where Elance kept its ads on. This result suggests that Elance did not need to purchase ads for its own term in order to block its closest competitor from poaching its users.

Beyond considering only advertising on trademarked terms, we assess the broader effects on oDesk's overall ad campaign of Elance shutting its ads off. We predicted that these effects would be large since we thought that Elance and oDesk's search ad campaigns were aggressively competing head-to-head, but instead we find relatively modest effects of smaller magnitudes than we expected, but in the expected directions.

We conclude our analysis with a comparison of the causal effect of Elance's search ads on its

[6]Navigational queries are extremely common. According to Mashable, in 2012, the top Google queries was for "Facebook" and the number 3 query was for "Google" itself. Queries for the name of a very well known website are generally considered navigational. See: `http://mashable.com/2012/12/07/top-google-searches-2012/`.

business with the typical correlation-based naive measures of ad efficacy. Without an experiment or some other source of exogenous variation in advertising, we could not tell if a potential customer who clicks an ad and registers on Elance would have otherwise registered absent the ad. With the data from a typical search ad campaign, we can track which registered users arrived at a website via a click on a search ad. We compare the number of registrants we can track to a click on a search ad (the naive estimate) with the number lost from turning the ads off (the causal estimate) and find that these estimates are significantly and substantively different. In particular, we find that the naive method overestimates the number of new users who registered due to search ads, and thus overestimates the value of these ads. The experimental estimate is 70% as large as the naive estimate. We conclude the analysis by estimating the impact of Elance's ads on oDesk's business, which we find to be small and insignificant.

The paper proceeds as follows. Section 1.2 provides additional legal and market context. Section 1.3 explains our experiment design. Section 1.4 presents our hypotheses. Section 1.5 describes the empirical methods we use to estimate our results. Section 1.6 shows the results of the experiment and we conclude in Section 1.7.

1.2 Legal and Market Context

1.2.1 Legal Context of Ads on Trademarked Search Terms

Search ads appear on search result pages for a variety of keywords, included both trademarked and non-trademarked terms.[7] Advertisements on trademarked keywords comprised 7% of Google's revenue as of April 2004,[8] the latest date for which this figure is publicly available. Revenues from trademarked keywords as a share of Google's total revenue are plausibly higher today, as Google has both permitted and encouraged more advertising on trademarked terms over the years.

[7]Trademarked terms are also known as brand terms in the advertising community. The conflicting meanings of "brand" in this context, as either trademarked keywords, or ads designed to promote a brand, are unfortunately confusing. We call ads triggered by trademarked keywords, "trademarked" ads for expositional purposes.

[8]Rosetta Stone Ltd. v. Google, Inc., 676 F.3d 144, 155-156 (4th Cir. 2012) (citing Joint Appendix, Vol. IX, Tab 41, Ex 6, "Google Three Ad Policy Changes" at p. 4264-4265). Rosetta Stone initially filed this case in 2009 and the parties settled in 2012.

In the U.S. prior to April 2004, Google allowed trademark holders to, upon request, block other advertisers from both advertising on their trademarked terms, and from including these terms in their ad text. Later in 2004, Google changed its policy to no longer allow trademark holders to block ads on their trademarked terms. Then in 2009, Google began to allow advertisers to include trademarked terms in their ad text under certain circumstances.[9] Additionally, Google has introduced and refined software tools, such as its Keyword Suggestion Tool, to suggest and aid in the discovery of relevant keywords, including trademarked terms, for advertisers to consider for their advertising campaigns.

Companies have complained to Google for allowing competitors to bid on search ads for their trademarked terms or to include these terms in search ad text. Several companies have sued Google over these issues, in both the United States and the European Union. In the U.S., in Rescuecom Corp v. Google and in Europe, in Louis Vuitton v. Google, the plaintiffs asserted that Google illegally allowed their competitors to bid on their trademarked terms.[10] Also in the U.S., in Rosetta Stone v. Google, Rosetta Stone asserted that Google illegally both allowed competitors to bid on Rosetta Stone's trademarked terms and to use these terms in their ad text.[11] Beyond direct concerns about possibly losing potential business to competing search advertisers, these companies argue in their cases that competing advertisers dilute their brands by confusing consumers, and in some cases, that these competitors are committing outright fraud, by misrepresenting themselves as the plaintiff's brands. The fact that advertisers spend billions of dollars per year specifically on search ads for trademarked terms explains why both Google has devoted resources to intricate trademark advertising policy changes and companies have initiated high-profile, costly lawsuits disputing search advertising on their trademarked terms. Bechtold and Tucker (2014) address the public policy issues surrounding the application of trademark law to search ads.

[9]Google's current AdWords Trademark Policy grants resellers and informational sites limited permission to use trademarked terms in their ad text. Google's current policy is available here: `https://support.google.com/adwordspolicy/answer/6118?hl=en`. Accessed August 8, 2014.

[10]Rescuecom Corp. v. Google, Inc., 562 F.3d 123, 126-127 (2nd Cir. 2009). Rescuecom initially filed this case in 2006 and dropped it in 2010. Court of Justice of the EU, Google France v. Louis Vuitton Malletier, Mar. 23, 2010, Joint Cases C-236/08 to C-238/08, ECR 2010, I-02417. Louis Vuitton initially filed this case in 2008, and the Court of Justice of the European Union ruled on it in 2010.

[11]Rosetta Stone Ltd. v. Google, Inc., 676 F.3d 144, 152 (4th Cir. 2012).

1.2.2 Market Context

After competing aggressively as the two largest online labor markets for years, Elance and oDesk merged during our experiment. Combined, freelancers completed $750 million of work on the two platforms in 2013.[12] A survey article provides an overview of how these markets work, the research being conducted on them, which countries are working and hiring, the types of work being performed and wage rates (Agrawal et al., 2013). They show a general pattern on oDesk of high-income country employers hiring low-income country freelancers, at hourly rates substantially higher than the minimum wages in the freelancer's country. Horton (2010) provides further background on the optimal design of these markets.

Elance and oDesk fought to bring new users to their platforms by purchasing search ads on a variety of terms, including both generic terms and each other's trademarked terms. A trademark owner has many possible reasons to care about the effects of competing advertisers bidding on their trademarked terms, beyond the direct effects on sales and search ad costs. Competing ads may dilute the value of a brand, and in some cases, have been misleading or fraudulent, as in the legal cases described in Section 1.2.1. In our setting, Elance cared specifically about potential business being lost to oDesk due to the the potential winner-take-all dynamics of their competing network-based online marketplaces. As such, Elance historically bid on its own name to keep oDesk from poaching potential customers via search advertising. Sayedi et al. (2014) propose a more complicated game to model the relationship between poaching in search advertising, and spending on traditional advertising, such as television and newspaper ads. In our setting, both Elance and oDesk primarily engaged in online advertising.

1.3 Experiment Design

The experiment design is fairly simple. During the experiment, Elance shut off all of its Google search ads in a randomly selected half of the Direct Marketing Areas (DMAs) in the United States

[12] `www.Elance-oDesk.com`. Accessed August 8, 2014.

for a period of 28 days, starting on March 11, 2014.[13] There are 210 DMAs, which subdivide the country into regions and they were originally designed for television-based advertising targeting. Advertisers can target Google ads geographically by DMA. The treated DMAs were selected completely at random, without stratification. After the experiment, Elance resumed bidding on search ads in the entire U.S., as it had before the experiment. Where relevant, we compare this experimental period to the 220 day pre-experimental period from Aug 1, 2013 to March 9, 2014. We use a longer pre-experiment period than the duration of the experiment to improve the precision of our estimated effect sizes. Prior to the experiment, Elance did not target or vary search advertising purchases within the U.S. geographically.

The experiment setup is similar to a an experimental study on eBay's search ad campaigns (Blake et al., 2014). Both experiments share one business goal of estimating amount of business generated by search ads, however, the novel setting of our experiment running during the Elance-oDesk merger process allows us to analyze the competitive effects of the experiment on oDesk's ad campaign and business.

Elance did not announce this experiment publicly to avoid strategic interference from other bidders. In particular, oDesk could have potentially disrupted the experiment by making different changes to its bidding behavior in the treatment and control DMAs during the experiment, if Elance told oDesk about the experiment, if oDesk discovered the experiment's effects independently, or if oDesk happened to make disruptive changes to its advertising campaigns during the experiment for reasons unrelated to the experiment. However, during the experiment, oDesk neither changed its bidding behavior overall nor did it specifically target its bids geographically within the U.S. The only changes oDesk made to its bids during the experiment in the U.S. applied to the entire U.S. These changes were minimal, and followed an overall bidding strategy that did not change during the experiment.

[13] Due to two independent problems, both unrelated to the experiment or its results, but which coincidentally occurred on consecutive days during it, there is a gap of seven days (which we call the "Omitted Period") without useful data during the experiment, so the effective experiment length is 21 days. The first problem is unexpected Elance bidding behavior (as described by the Elance marketing department) that occurred for three days resulting in substantially fewer ad impressions those days, and the second is a denial of service attack against Elance which resulted in four days of lost or unusable data. Days in which either of these two problems occurred, along with the day before the start of the experiment (during which Elance tested out the experiment for part of the day) are omitted from our analysis.

The two companies concluded the process of merging coincidentally during the experiment, one week prior to the end of the experiment. Prior to the conclusion of the merger, the companies needed to operate as separate, competing entities, and as such, Elance did not inform oDesk about the experiment. In preparation for the anticipated conclusion of the merger, neither Elance nor oDesk made any substantial changes to their bidding during the experiment period prior to the merger (aside from Elance running this experiment). Following the conclusion of the merger, no major changes were made to oDesk's bidding strategy until after the end of the experiment period, both to facilitate completion of the experiment and to allow the new company enough time to formulate an updated search advertising strategy. We return to the subject of a competitor's best response to an advertiser dropping out of an ad auction in Section 1.7, our conclusion.

The experiment's duration was limited by business concerns. Simulation-based a priori power analysis suggested that we run the experiment for a minimum of two weeks, a duration we expected would yield somewhat imprecise, but usable results to assess the overall performance of Elance's ad campaign. As noted across a variety of ad campaigns by Lewis and Rao (2013), attaining a large enough sample size to achieve the statistical power necessary to evaluate a campaign can be a significant challenge. In the case of our experiment, ideally, we would be able to randomize at the individual level, rather than the DMA level, to gain more power, however, doing so is not possible with currently available ad targeting and tracking options.

1.4 Hypotheses Development

1.4.1 Effects on oDesk's Search Ad Campaign

Google auctions its ads using the Generalized Second-Price (GSP) mechanism, described in detail in Appendix A1. Table 1.2 describes the metrics which summarize the ad campaign, and our prediction of the effect the experiment would have on oDesk's campaign overall, and specifically on the "Elance" term. Based on the features of the GSP mechanism, we expected that oDesk's cost per click (CPC) would go down whenever Elance's ads would otherwise have occupied the ad position one below (i.e. worse) than oDesk's, and otherwise oDesk's CPC would remain unchanged.

We thus expected oDesk's CPC to go down overall. Similarly, oDesk's position would decrease (i.e. improve) by exactly one if Elance's ad would outrank oDesk's, and otherwise would remain unchanged. We thus expected oDesk's position to improve (i.e. decrease) overall. oDesk's position improvement would equal the share of auctions in which its ad and Elance's ad were displayed, and Elance's ad's position was better.

Hence, we know the expected direction of the effects of the experiment on oDesk's CPC and position based on the search ad auction's features, but the magnitudes of these effect sizes depend on how aggressively oDesk and Elance were competing with each other with their search ad campaigns. If Elance and oDesk had both bid in all of the same search ad auctions and bid equally aggressively so that their quality score adjusted bids (see Appendix A1) were generally close, then oDesk's position would have improved by approximately 0.5 and its CPC would have decreased substantially (as long as, in the cases where Elance was one position worse than oDesk, the next highest bidder's quality adjusted bid was substantially lower than Elance's). If instead the two companies did not bid in any of the same auctions, then neither oDesk's CPC nor position would change. If there was overlap, but oDesk outranked Elance with all of its ads, then position would not change at all, but CPC would still potentially go down. Based on internal Elance market research, we believed that oDesk and Elance were fierce competitors for search ads, so we expected the experiment to decrease oDesk's CPC substantially. Since the two companies did not target the exact same set of keywords, we thus expected position to improve by somewhat less than 0.5, but still by a substantial amount. For the term "Elance", since Elance was in position 1 and oDesk was in position 2 generally, we expected oDesk's position to decrease by 1 (becoming 1), but its CPC to remain unchanged.

oDesk's ad impressions count would go up in cases where the Elance ad would have been the worst-ranked ad shown, and oDesk's ad was the highest ranked ad that was not shown. This specific situation is rare overall, and was non-existent for the term "Elance," so we expected a very small overall increase in impressions overall, but no increase for the term "Elance." Clicks generally increase with more impressions and better position, so we expected clicks on oDesk ads to increase overall, primarily due their improved position, and to increase for the term "Elance" entirely due to the improvement in position from 2 to 1. We further expected oDesk to get more clicks due to the absence of Elance's ad, independently of the effects of clicks due to position and impressions. Ads in position 1 generally receive substantially more clicks than ads in position 2 (Jansen et al., 2013), so we expected a particularly large increase in clicks on oDesk's ads on the term "Elance," a key prediction we will return to in section 1.6.1. We made no prediction about the overall cost of oDesk's ad campaign. We expected more clicks (increasing cost), but lower CPC (reducing cost), and had no general expectation about the relative size of these two effects, but we did expect cost on the term "Elance" to increase substantially, via an increase in clicks, but no change in CPC.

1.4.2 Effects on Elance and oDesk Business

Online advertisers typically track incoming visitors to their sites to help them assess the efficacy of their ad campaigns. Typically, online advertising software, including Google's will show an advertiser the number of new users (or other metric the advertiser chooses) that clicked on an ad. Advertisers typically use this observational, correlation-based measure as a naive estimate of the amount of new business they gain from an ad campaign. It is naive because it assumes a possibly unrealistic counterfactual, that absent the ads, none of the users who clicks on them would have signed-up for the advertiser's site. Previous research (Lewis et al., 2011) conducted randomly controlled trials in advertising across a variety of online settings and demonstrated by comparing observational and experimental estimates that observational methods can drastically overestimate the efficacy of online advertisements. For example, their first experiment (Section 2 in their paper) showed that the naive estimate of an advertising campaign was over 200 times as large as the causal point estimate, and that this naive estimate decreased only slightly when adding in relevant controls. Blake et al. (2014) found similar overall results when experimentally analyzing eBay's search ads. Their Table 3 shows a small, statistically insignificant causal effect of their ads, compared to a naive estimate on the order of 100 times as large (depending on which controls are used).

Both Elance and oDesk primarily used search ads to attract new employers to their services, so the number of new employer registrations is the metric we use to quantify the effects of Elance's ads on both businesses. Table 1.3 shows our predictions of the experiment's effect on new employer registrations for both companies. Our experiment provided us with an estimate of the percentage of Elance's new employer registrations resulting from search ads (i.e. that we tracked to a click on a search ad) which we should attribute to the choice to run the ad. Adver This percentage is equal to 100% only if none of the new registrations would have occurred absent the ads, meaning that a potential new employer would not have found Elance via an organic search result or some other method of discovery. We expected some of this substitution to other ways of discovering or deciding to use Elance, but not a tremendous amount since Elance was not a generally well known brand (we will return to this point in the conclusion). Overall, we thought that Elance would lose a substantial amount of new business by turning its ads off. Elance dropping out of the ad auction would have to at least weakly help oDesk's business by increasing its ads' exposure and number of clicks. Based on our predicted effects on oDesk's overall ad campaign as described in Table 1.2 and a pre-merger estimate that search ads accounted for a large, but non-majority share of oDesk's new employer registrations, we expected a slight increase in this metric for oDesk.

1.5 Empirical Methods

To account for pre-experiment differences across the different DMAs, we evaluate all results using a difference-in-difference approach via the following fixed-effects regression:

$$Y_{it} = f(\beta_1 * \text{AdsOff}_{it} + \delta_t + \gamma_i + \epsilon_{it}) \tag{1.1}$$

i indexes the 210 different DMAs, t indexes time periods and $f(\cdot)$ is a function. We aggregate the results into two time periods: "before" and "during" the experiment. Y_{it} is the outcome variable and β_1 is the coefficient of interest. AdsOff_{it} is an indicator variable which is equal to 1 for observations of treatment DMAs during the experimental period. δ_t and γ_i are the time and DMA fixed effects, respectively, and ϵ_{it} is the error term. In all results, we cluster standard errors at the DMA level.

We wish to model the effect of turning Elance's ads off as causing a percentage change in some outcome variables, including the number of new registrations per DMA, rather than a change in levels. Different DMAs represent different populations, both overall, and of pre-treatment Elance and oDesk employer registrations. The amount spent per DMA on advertising pre-treatment is roughly in proportion to each of these population. Thus, for example, if we consider two DMAs, A and B, where A represents ten times the population of B, we would expect the search ads to cost ten times as much in A as B and to bring in ten times as many employers, but to bring in a similar percentage of total new employers within each DMA.

For variables where we expect and wish to estimate a percentage change, we model $f(x) = exp(x)$ and estimate the results using the Poisson quasi-maximum likelihood estimator (QMLE).[14] We prefer this estimator to taking the log of Y_{it} and estimating via OLS because some of our outcome observations are equal to 0.[15] For other variables where we expect a linear change, primarily the ad position, we simply model $f(x) = x$, and estimate via OLS.

1.6 Results

We first present and discuss the raw data, foreshadowing the formal experimental results which we turn to afterwards. Table 1.4 summarizes before and during experiment data, in both experiment cells (the treatment and control group DMAs). Figures 1.2, 1.3 and 1.4 show time series plots of the experiment data for oDesk's ad campaign for the term "Elance" (Figure 1.2) and overall ad campaign (Figure 1.3), and oDesk and Elance's business metrics (Figure 1.4).

Table 1.4 shows that oDesk's ads on the term Elance had a large change in position (-0.96) and substantial changes in clicks and cost in the "difference-in-difference" cells of the table. Other variables in oDesk's campaign on the term Elance, as well as all variables in oDesk's entire campaign, did not change substantially. Elance's user registrations were down substantially in the

[14]Silva and Tenreyro (2006) and Wooldridge (2002) describe and motivate the use of this estimator. This estimator does not assume that $Var(Y|X) = E(Y|X)$, as the name Poisson might misleadingly suggest, for consistency or asymptotic normality, and it has nice efficiency and robustness properties (Wooldridge, 2002).

[15]In practice, for some outcome variables there are few observations equal to zero, and in these cases, we get similar results when we estimate using OLS, and either use $log(Y_{it} + 1)$ as our outcome variable, or drop observations where $Y_{it} = 0$.

treated DMAs, but oDesk's were not.

In Figures 1.2 and 1.3, we see that the search ad metrics varied over time and that oDesk's overall ad campaign followed somewhat different patterns from its ads on the term Elance. In Figure 1.2, oDesk's position for the term Elance was relatively consistent before the experiment, yet oDesk's position varied much more overall as seen in Figure 1.3, and has some distinct trends over time. Similarly, impression counts fluctuated more overall than for the term Elance. Clicks and cost fluctuated substantially both overall and for the term Elance. Like position, cost per click (CPC) follows some time trends overall as well. Overall, there were only very slight differences in these metrics between the two experimental cells in the pre-experiment period.

Before the experiment, oDesk did not bid on the term Elance for a short period, as seen in all of the panels of the Figure 1.2. Impressions for the term Elance increased modestly several months before the start of the experiment, and substantially about a week before, remaining higher post-experiment than before. Impression counts could have varied due to a variety of changes in the following areas: number of searches for this term, oDesk's bidding behavior, other advertisers' bidding behavior and Google's auction algorithm.

Figure 1.4 shows that Elance registered almost the same number of new clients pre-experiment in the treatment and control cells, yet fewer during the experiment in the treatment group. oDesk registered fewer clients in the treatment group both before and during the experiment.

1.6.1 The Experiment's Effects on oDesk's Advertising Campaign

We first analyze the effects of the experiment on oDesk's Google ad campaign for searches on the keyword "Elance" (Table 1.5), then for oDesk's overall Google ad campaign (Table 1.6), by estimating Equation 1.1 using the different ad outcome metrics described in Table 1.2. In all cases, the treatment group is the set of DMAs where Elance turned its ads off, and in the control group, they remained on. The results confirmed some of our predictions, but refuted others.

First, as expected, oDesk's ad position for the keyword "Elance" improved by almost exactly 1, and impressions remained almost exactly the same (Table 1.5, columns 1 and 2), indicating that the treatment worked as expected for these terms. But, surprisingly, even with both this improvement in position and with Elance's ad out of the way, clicks did not increase (column 3), and in fact, the point estimate is negative (-18%), though not statistically significant. If oDesk had instead captured all of the search ad clicks Elance lost by not running its ads, oDesk would have received approximately 10,000% more clicks (a coefficient of approximately 100). We reject the hypothesis that oDesk even doubled their clicks, an effect $\frac{1}{100}$ as large (a coefficient of approximately 1) and a result which would have still yielded an insubstantial amount of new business for oDesk. In terms of the simple model represented by Table 1.1, where Table 1.1a represents a simple Prisoner's Dilemma game where bidding on one's own trademarked terms blocks competitors from getting searchers to click their ads, and Table 1.1b represents an alternative game where such bidding is less effective, these results show that Table 1.1b better models the competition between Elance and oDesk for ads on the term "Elance". oDesk's changes in cost and CPC are not statistically significant (columns 4 and 5), but are not precisely estimated.

The effect on oDesk's overall ad campaign was smaller than expected. Position did decrease as expected, but the estimated, statistically significant effect (Table 1.6, column 1) implies that only 5.5% of searches which triggered an oDesk ad contained a better positioned Elance ad. Clicks increased as expected (columns 3), although the point estimate of 2.8% is not statistically significant. Likewise, CPC decreased by a statistically significant 7.0% (column 5). If these effects were larger, especially the CPC estimate, they would have suggested that the merged company was harming itself by bidding against itself, and thus should bid less aggressively in one or both of its campaigns, or make other changes to avoid competing with itself.

1.6.2 The Experiment's Effects on Elance's and oDesk's Businesses

We next analyze the effect of shutting down Elance's Google ads on Elance's and oDesk's businesses. Table 1.7 shows the business impact of the experiment as estimated by using Equation 1.1, on the new user metrics described in Table 1.3. The point estimate in column 1 suggests that Elance lost approximately 23% of its new employers by turning its search ads off.[16]

We wish to test two hypotheses about the efficacy of these ads for Elance. The first is that

[16]The experiment did not introduce separate exogenous variation in advertising on trademarked and non-trademarked terms, so we are unable to identify separate impacts of advertising for these two groups of terms.

the ads were completely ineffective, and that any relationship between the ads and new employer registrations is entirely non-causal. This is a test that $\hat{\beta}_1 = 0$ in Table 1.7, column 1. We reject this hypothesis at the conventional 0.05 level using a two-sided t-test, with a p-value of < 0.001, as the estimate is both statistically and substantively significantly higher than zero. The second hypothesis is that each employer who registered after clicking a search ad was causally induced to register because of the ad. In other words, this hypothesis states that without search ads Elance should have lost the proportion of new employers tracked to a search ad click in the control group. Based on the proportion of new registrations Elance tracked to a search ad click in the control group, this is a test that $\hat{\beta}_1 = -0.338$ in Table 1.7, column 1. This assumption, that each click tracked through to a registration should be credited to the ad campaign, is a standard one marketers would use to assess the performance of a search ad campaign, if they were not concerned about causality. We also reject this hypothesis at the conventional 0.05 level using a two-sided t-test, with a p-value of 0.0002.

We thus estimate that the causal impact of these ads is about 70% of the naive estimate. Thus, this experiment provides evidence that these search ads were less than 100% as effective as the naive estimate would suggest. However, the gap between the experimental and naive estimates in this experiment (the naive estimate is 1.4X the causal estimate) is much smaller than the ones implied by Blake et al. (2014) for search ads (a 100X difference) and Lewis et al. (2011) for online display ads (a 200X difference), suggesting that the naive methods for assessing online advertising campaign effectiveness are less biased for our campaign than for the campaigns in these experiments. We suspect that the primary reason for this difference is because Elance is a less generally well-known brand than the brands in these other studies. However, the fact that this gap is still substantial suggests that even in Elance's case, optimal bidding must take it into account. Following the completion of the experiment, the merged company used this result as an input to its updated bidding strategy, by bidding lower than the estimated optimal level implied by the naive estimates.

Finally, we also assess the impact on oDesk's business. We fail to reject the null hypothesis that Elance's ad campaign had no effect on oDesk, that is, that $\hat{\beta}_1 = 0$ in Table 1.7, column 2. The point estimate suggests that oDesk registered approximately 0.9% more employers as a result of Elance turning its ads off. This small estimated impact of Elance's search ads on oDesk's business is unsurprising considering the modest impact of Elance's search ads on oDesk's search ads that we estimated in Section 1.6.1.

1.7 Conclusion

Our results show that oDesk did not gain a significant amount of the search ad traffic Elance lost when it stopped bidding on its own search term. Most experts would not have predicted this result, and they instead claim that companies must bid on their own terms to prevent competitors from reaching their customers. This discrepancy raises the question of whether our results generalize to other advertisers bidding on their own trademarked search keywords. If it does, then many advertisers are needlessly spending money defending their trademarked terms by bidding on them. We suspect this result does in fact generalize in many cases, because Elance and oDesk both behaved like typical advertisers on Google and there was nothing particularly unusual about their competition over search ads for each other's trademarked terms. Both companies were large search advertisers during the experiment, but were far from being the largest. Like the vast majority of search advertisers, neither had conducted a randomized controlled trial with their search ads prior to this experiment. Both companies bid on a variety of keywords, not just trademarked terms. For some keywords, they faced intense competition from each other and other advertisers, while for others they did not. There is nothing to suggest that Elance and oDesk's competition over search ads for the term "Elance" was in any substantial way different from search ad competition between other businesses over their trademarked search terms.

In our experimental setting, oDesk did not respond to Elance stopping some of their ads as described in Section 1.3. In a more general setting, a competitor might respond to a business stopping some or all of their search ad purchases by bidding more or less aggressively, either overall, or only on some keywords, especially over a longer time horizon than our short experiment period.[17]

[17] Note that the GSP auction is not incentive compatible (Edelman et al., 2007) and see Appendix A1 for further details on the GSP mechanism.

Suppose Elance and oDesk remained competitors, and that Elance stopped bidding on search ads. For the term “Elance,” oDesk should not have responded, but rather would have passively moved from position 2 to position 1, and unhappily found out that their traffic did not increase substantially from this term. For its overall ad campaign terms, oDesk’s best move to reoptimize depends on the other remaining bidders, so we we are unable to predict the best response to Elance dropping out of the auction, but because of the relatively small impact Elance’s participation had on oDesk’s campaign, we suspect oDesk’s optimal bid changes in this scenario would be small. If oDesk bid to spend a fixed marketing budget, which is a common, though not universal practice among search advertisers, then oDesk’s lower CPC would allow it to increase its bids and acquire more clicks. Of course, if Elance were to drop out of the auction, it would lose the substantial amount of new business these ads generated for it regardless of exactly how oDesk would change its bids in response.

The results in Blake et al. (2014) suggest that eBay’s search ads were not very effective in many cases. The authors attribute this result to eBay being a well-known brand. Users who clicked on a search ad and subsequently made a purchase on eBay would likely have made their purchase if not shown the ad because they already knew about eBay. We share this view, and because Elance is a far smaller company than eBay and is substantially less well known, we did not necessarily expect their results to generalize to Elance’s overall search ad campaign efficacy. Rather, our results complement theirs and together provide evidence consistent with the view that the gap between the naive and causal estimates of a company’s search ad campaign effectiveness increases in the company’s size. Following our experiment, we now believe that employers clicking on Elance’s search ads generally would not have signed up without the ads.

Table 1.1: Game models of the search ad auction when a trademark owner bids against a competitor, with best responses in bold.

		Competitor: Bid	Competitor: Don't
Owner	Bid	**-3**,**-1**	**3**,-4
	Don't	-4,**3**	0,0

(a)
Stylized payoff matrix for the search ad bidding game for trademarked terms if it is truly a Prisoner's Dilemma. In this case, the trademark owner should defend its keyword by bidding.

		Competitor: Bid	Competitor: Don't
Owner	Bid	-3,**-1**	**3**,-4
	Don't	**-2**,**1**	0,0

(b)
Stylized payoff matrix for the search ad bidding game for trademarked terms if the competitor does not gain as much business as in 1.1a when the owner does not bid. In this case, the trademark owner should not bid.

Table 1.2: Description of advertising metrics and predicted effects of the experiment on oDesk's Google ad campaign.

Outcome	**Prediction (Ads using the "Elance" term)**	**Prediction (All terms)**	**Description**
Cost per click (CPC)	Remain the same	Decrease substantially	Total cost divided by total clicks.
Position	Decrease by ≈ 1	Decrease by < 1	The average position of the ad on the search results page. 1 indicates the top position on the page, which is most likely to get clicked.
Impressions	Remain the same	Increase slightly	Count of total times an ad is shown.
Clicks	Increase substantially	Increase substantially	Count of total times an ad is clicked on.
Cost	Increase substantially	Unsure	Total cost of the ad campaign.

This table shows the predicted effects of turning Elance's search ads off on oDesk's ads using the "Elance" trademarked keyword, and all oDesk ads, for each of five search ad metrics.

Table 1.3: Description of Elance and oDesk business metrics and predicted effects of the experiment on them.

Outcome	**Prediction**	**Description**
Elance Employer Registrations	Decrease substantially, but by somewhat less than number of new employers who clicked on a search ad	Number of new employers who register on Elance
oDesk Employer Registrations	Increase slightly	Number of new employers who register on oDesk

This table shows the predicted impact of turning Elance's search ads off on Elance and oDesk's business outcomes.

Table 1.4: Mean variable values before and during the experiment, by experiment cell, with differences and difference-in-differences

Variable	Control	Treatment	Difference
oDesk ads on "Elance" term			
Position, before experiment	1.97	1.98	0.01
Position, during experiment	2.11	1.17	−0.94
Difference	0.14	−0.81	−0.96
Impressions, before experiment[1,2]	1.00	0.97	−0.03
Impressions, during experiment[1,2]	2.26	2.21	−0.05
Difference[1,2]	1.26	1.24	−0.02
Clicks, before experiment[1,2]	1.00	1.16	0.16
Clicks, during experiment[1,2]	1.64	1.59	−0.05
Difference[1,2]	0.64	0.43	−0.22
Cost, before experiment[1,2]	1.00	1.06	0.06
Cost, during experiment[1,2]	2.16	1.93	−0.23
Difference[1,2]	1.16	0.87	−0.29
Cost per click, before experiment[1]	1.00	0.91	−0.09
Cost per click, during experiment[1]	1.31	1.21	−0.10
Difference[1]	0.31	0.30	−0.01
oDesk ads entire campaign			
Position, before experiment	3.42	3.45	0.02
Position, during experiment	3.71	3.67	−0.03
Difference	0.28	0.23	−0.06
Impressions, before experiment[1,2]	1.00	0.99	−0.01
Impressions, during experiment[1,2]	0.81	0.79	−0.02
Difference[1,2]	−0.19	−0.20	−0.01
Clicks, before experiment[1,2]	1.00	1.00	−0.00
Clicks, during experiment[1,2]	0.98	1.00	0.02
Difference[1,2]	−0.02	0.01	0.03
Cost, before experiment[1,2]	1.00	1.00	−0.00
Cost, during experiment[1,2]	1.17	1.16	−0.01
Difference[1,2]	0.17	0.16	−0.01
Cost per click, before experiment[1]	1.00	1.00	0.00
Cost per click, during experiment[1]	1.19	1.15	−0.04
Difference[1,2]	0.19	0.15	−0.04
Elance and oDesk new user registrations			
Elance registrations, before experiment[1,2]	1.00	0.99	−0.01
Elance registrations, during experiment[1,2]	1.02	0.80	−0.22
Difference[1,2]	0.02	−0.19	−0.21
oDesk registrations, before experiment[1,2]	1.00	0.89	−0.11
oDesk registrations, during experiment[1,2]	1.15	1.03	−0.12
Difference[1,2]	0.15	0.14	−0.01

Some variables are normalized to protect proprietary data.
1 Variable normalized to equal 1 in the control group before the experiment.
2 Variable normalized per day.

Table 1.5: oDesk ads on “Elance” term estimates

	(1)	(2)	(3)	(4)	(5)
Y_{it}:	Position	Impressions	Clicks	Cost	Cost Per Click
Method:	OLS	QMLE	QMLE	QMLE	QMLE
AdsOff_{it}	-0.938***	0.009	-0.184	-0.171	0.230
	(0.033)	(0.059)	(0.135)	(0.125)	(0.145)
N	408	420	420	420	241

Each column shows the estimated impact of Elance shutting down its Google ads on different aspects of oDesk's Google ad campaign for the term “Elance”. All estimates include DMA and time-period fixed effects and are estimated using Equation 1.1. Standard errors are clustered at the DMA level and are in parentheses. Columns (1) and (5) contain fewer than 420 samples because position and CPC are only defined if there are any impressions and clicks, respectively, in a time period-DMA observation.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.6: oDesk ads entire campaign estimates

Y_{it}:	Position	Impressions	Clicks	Cost	Cost Per Click
Method:	OLS	QMLE	QMLE	QMLE	QMLE
AdsOff_{it}	-0.055**	-0.018	0.028	-0.007	-0.070*
	(0.017)	(0.019)	(0.021)	(0.027)	(0.035)
N	408	420	420	420	408

Each column shows the estimated impact of Elance shutting down its Google ads on different aspects of oDesk's entire Google ad campaign. All estimates include DMA and time-period fixed effects and are estimated using Equation 1.1. Standard errors are clustered at the DMA level and are in parentheses. Columns (1) and (5) contain fewer than 420 samples because position and CPC are only defined if there are any impressions and clicks, respectively, in a time period-DMA observation.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

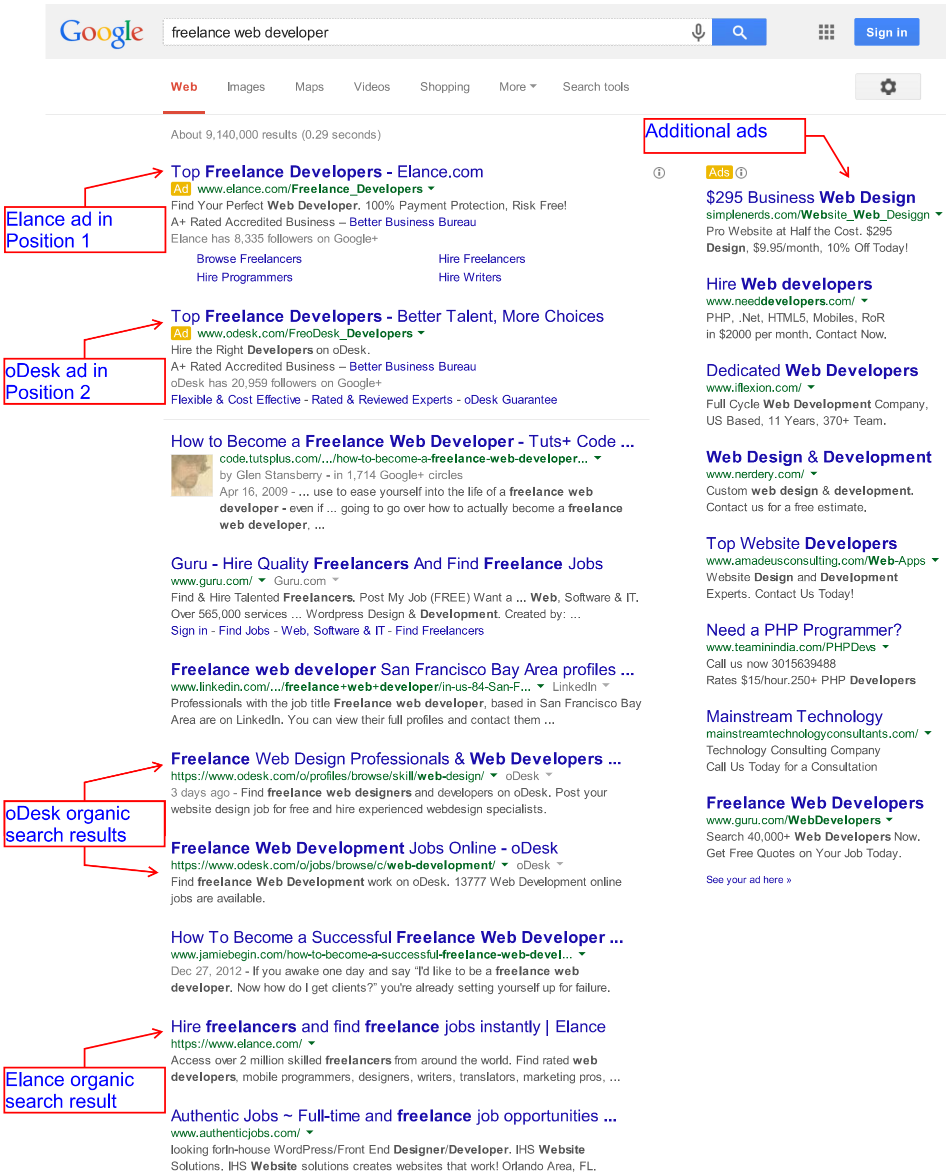
Table 1.7: New user registrations estimates

	(1)	(2)
Y_{it}:	Elance Registrations	oDesk Registrations
Method:	QMLE	QMLE
AdsOff_{it}	-0.229***	0.009
	(0.030)	(0.036)
N	420	420

Each column shows the estimated impact of Elance shutting down its Google ads on either Elance or oDesk's new client registrations. All estimates include DMA and time-period fixed effects and are estimated using Equation 1.1. Standard errors are clustered at the DMA level and are in parentheses.

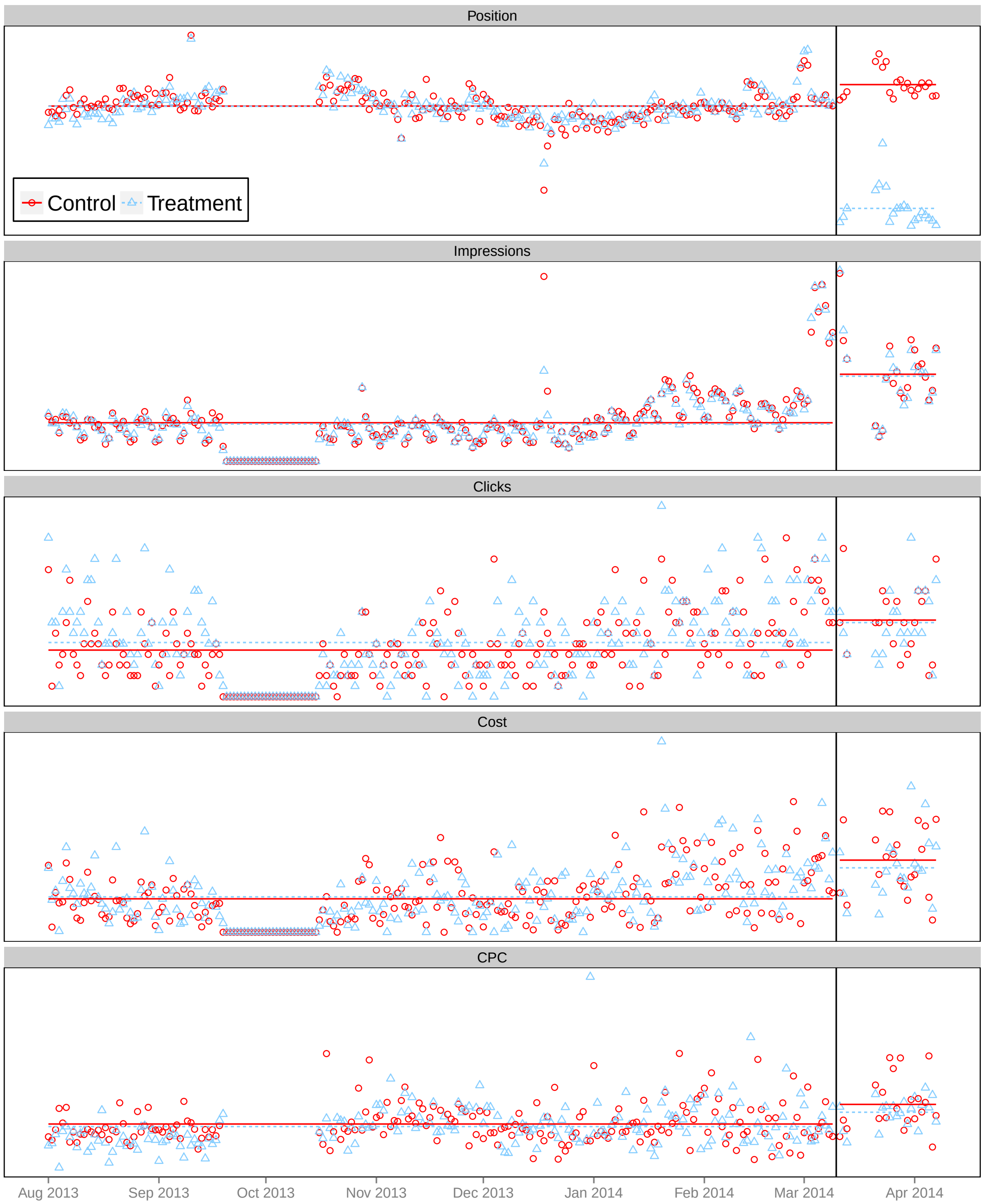
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1.1: Google search results including oDesk and Elance ads and organic results



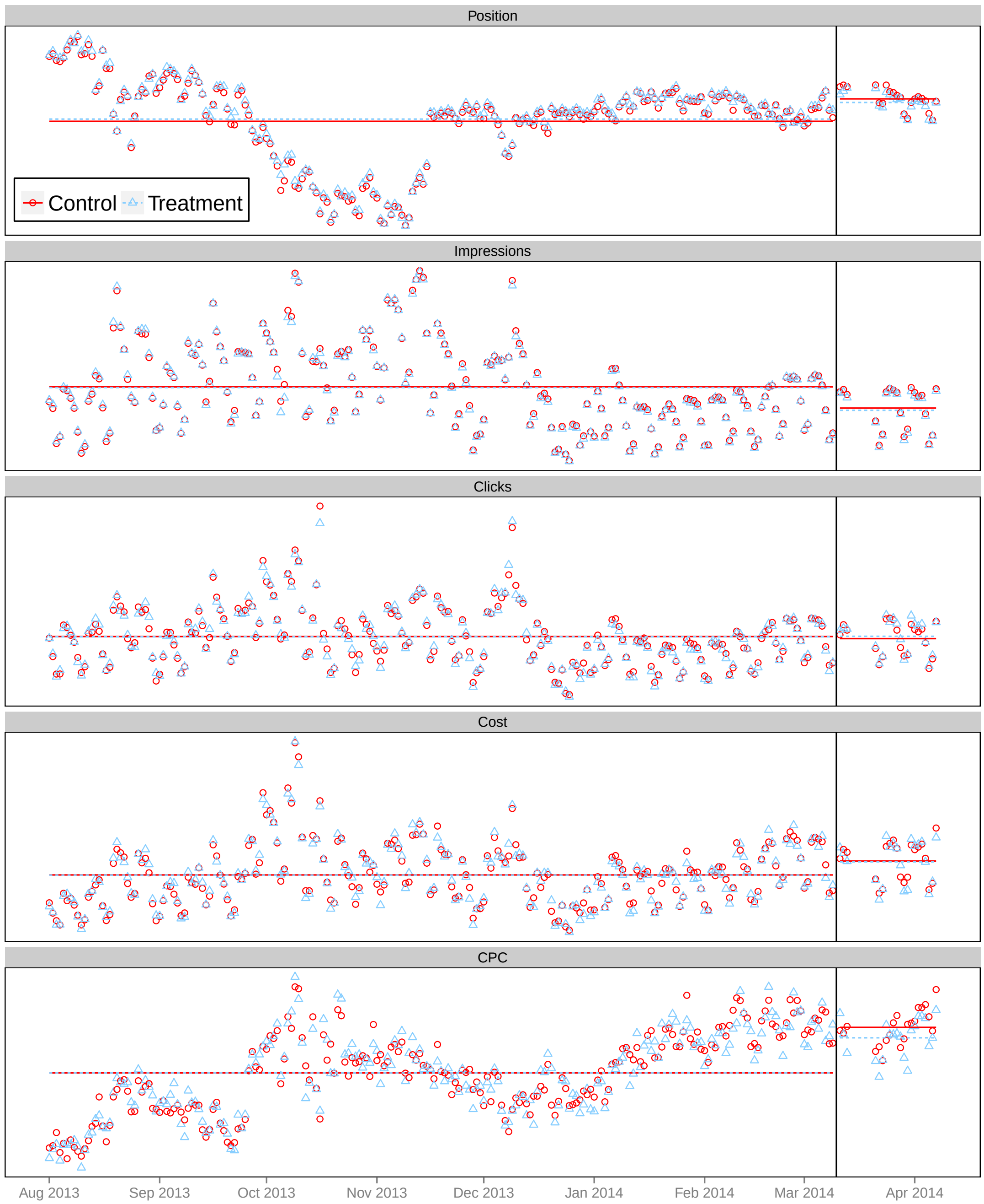
This annotated screenshot shows an illustrative Google search results page from a search query on which Elance and oDesk bid on search ads. It shows an Elance ad in position 1 and an oDesk ad in position 2 in the top ad block, along with 8 other ads from other businesses in the right ad block. Elance appear as the seventh organic search result, and oDesk appears as the fourth and fifth organic search results. This search was conducted on June 6, 2014.

Figure 1.2: oDesk search ad campaign for the term “Elance”



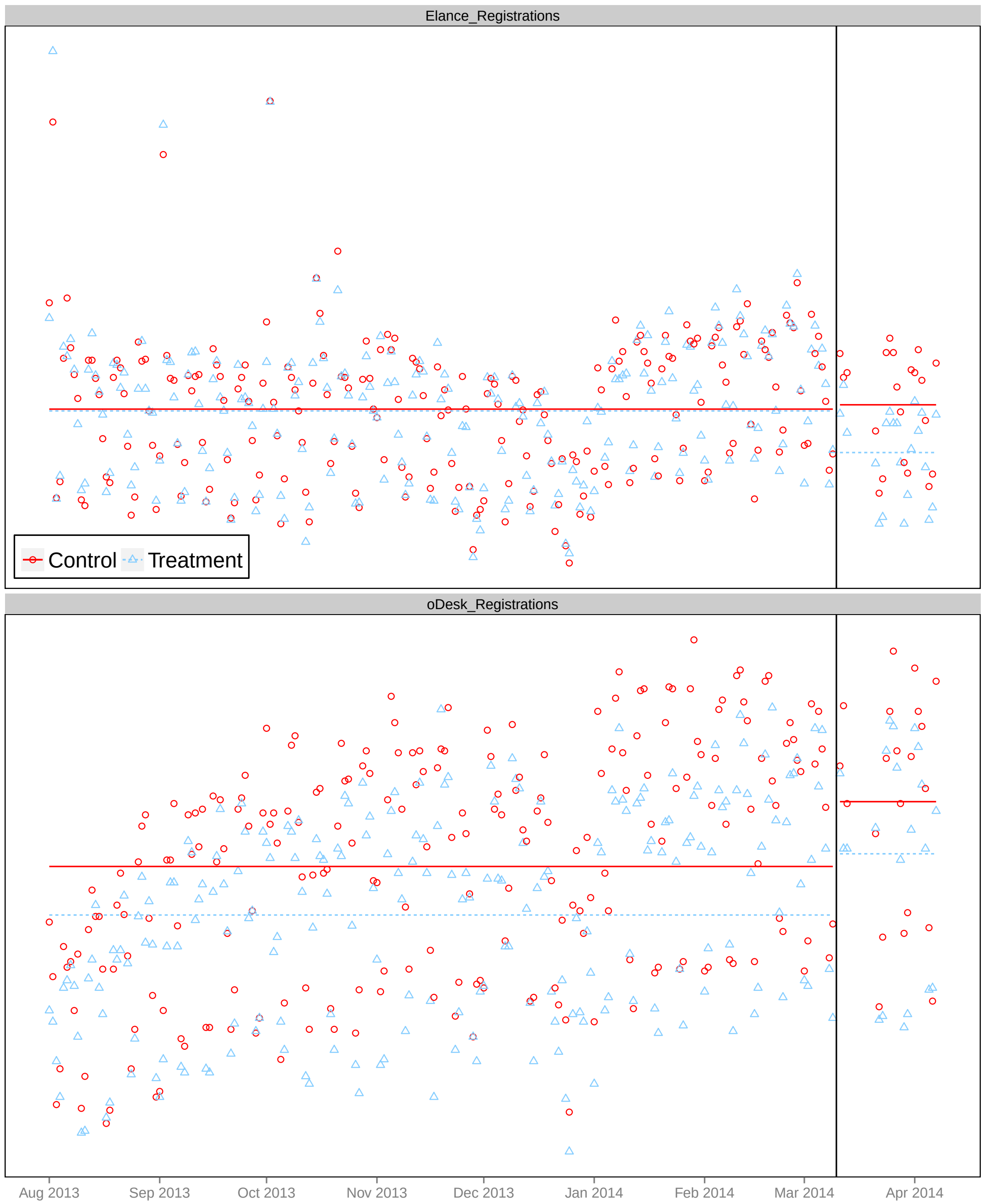
This Figure shows daily search ad metrics for oDesk’s search ad campaign for the term “Elance”. The vertical line in each plot separates the before and during experiment periods. The horizontal lines are the means for the treatment and control groups, during each of the two periods. Randomization was done at the DMA level. Observations are at the DMA-day level and each data point represented in the figure is an aggregation to the experiment cell-day level. The vertical axes are unlabeled to protect proprietary data. Data are omitted for the seven day “Omitted Period” during the experiment period, as described in Section 1.3.

Figure 1.3: oDesk search ad campaign, overall



This Figure shows daily search ad metrics for oDesk's overall search ad campaign. The vertical line in each plot separates the before and during experiment periods. The horizontal lines are the means for the treatment and control groups, during each of the two periods. Randomization was done at the DMA level. Observations are at the DMA-day level and each data point represented in the figure is an aggregation to the experiment cell-day level. The vertical axes are unlabeled to protect proprietary data. Data are omitted for the seven day "Omitted Period" during the experiment period, as described in Section 1.3.

Figure 1.4: Elance and oDesk new client registrations



This Figure shows daily new client registrations for Elance and oDesk, which both aimed to increase via search advertising. The vertical line in each plot separates the before and during experiment periods. The horizontal lines are the means for the treatment and control groups, during each of the two periods. Randomization was done at the DMA level. Observations are at the DMA-day level and each data point represented in the figure is an aggregation to the experiment cell-day level. The vertical axes are unlabeled to protect proprietary data. Data are omitted for the seven day "Omitted Period" during the experiment period, as described in Section 1.3.

A1 Description of Search Ad Auctions

This Appendix provides the relevant background material describing search ad auctions to give context to the experiment presented in this paper. For further details, please see Varian (2007) and Edelman et al. (2007), two papers which provide a more general overview and analysis of search ad auctions.

When users search Google, it generates two separate sets of ranked results related to the search term: organic search results and paid ads. Figure 1.1 shows an example search engine results page which includes both search results and paid ads. Other popular search engines, such as Bing, work similarly, and face similar constraints and objectives. We will describe the system which determines which ads get displayed, in which positions and at what cost to the advertisers.

Google sells its search ads via "generalized second-price" (GSP) auctions. These algorithmic auctions happen in real-time, nearly instantly, triggered by each search. Google's ad inventory consists of potential ad positions in which to show an ad impression, up to some maximum number per page. This inventory is highly heterogeneous, as advertisers target their ad copy and their bids, which they submit in advance, to specific search terms.

We can think of Google's short-term goal as maximizing expected revenue from each search, and its long-term goal as maximizing expected discounted revenue from future searches. Although Google's unit of inventory is an impression, advertisers generally submit bids not on impressions, but on clicks on the search terms on which they wish to advertise. These bids are known as cost per click (CPC) bids. Google computes a bid-modifying quality score for each advertisement in an auction as a function of various quality metrics (Varian, 2007), including Google's estimate of the ad's click-through rate (for a given position), which is the percent of users who see the ad that click on it. Search engines, including Google, generally do not make public their exact methods for scoring an ad. In the short-term, Google maximizes revenue by maximizing the cost per impression to advertisers, which is equal to the cost per click, multiplied by the click-through rate. However, seemingly paradoxically, Google generally displays no ads for some of the most popular search queries, searches for the names of the most popular websites, such as "Google" and "Facebook". These queries are generally navigational, meaning that users are searching for the names of these sites in order to navigate to them, not to discover other search results and more information, thus a relatively low proportion of users would likely click on ads on results pages for these terms. In the short term, ads on these terms would still generate positive revenue and it would be optimal for Google to show them. But, showing poor quality ads may discourage users from paying attention to all ads in the future, harming future revenue, which is one reason why some searches yield zero ads or fewer than the maximum number of ads (another more obvious reason is if there are few or no bidders for a particular search term).

Ads are positioned by the rankings of their scores. There is generally a reserve price in an ads score and/or bid. When a user clicks on an ad in position i, the GSP mechanism determines that the advertiser pays the minimum amount (or slightly more) that would keep their ad's score just above the ad in position $i + 1$. That is, the advertiser pays approximately the following per click:

$$\text{CPC}_i = \frac{\text{bid}_{i+1} * \text{score}_{i+1}}{\text{score}_i} \tag{A1}$$

See Figure 1.1 for an example of the results of one instance of this auction process.

Advertisers seek to maximize their surplus from the search ad auction as a function of their bid (including the choice to not participate by not bidding). As discussed in Varian (2007) and Edelman et al. (2007), the GSP mechanism differs from the Vickrey-Clarke-Groves (VCG) mechanism in that it has no dominant strategy equilibrium, and truth-telling is not an equilibrium. Rather, advertisers can have an incentive to bid below their valuation, because in some cases doing so results in a less prominent, but less expensive ad position which yields higher surplus to the advertiser. The GSP mechanism is thus subject to strategic manipulation, even when advertisers know their true valuations. In practice, advertisers do not know their true valuations and estimating them is difficult, adding additional complexity to their optimal bidding problem.

Edelman et al. (2007) document the early history of search ad auctions, and the GSP mechanism's predecessors which led to it, including a generalization of the first-price auction mechanism introduced by Overture (now part of Yahoo) in 1997 where advertisers paid their bid for each click. This mechanism led to inefficient outcomes in which advertisers developed bidding robot software to rapidly manipulate their bids (to attempt to just barely outbid competitors). Varian and Harris

(2014) recounts the early history of search ad auctions at Google, which first started selling search ads via an auction in 2002, using the GSP mechanism (initially developed primarily by computer engineer Eric Veach), to address the problems with the generalized first-price mechanism. Recognizing the advantages of a truthful mechanism, Google later considered switching from GSP to VCG, but did not for several reasons, one of which is that advertisers would have to raise their bids to re-optimize their campaigns if this switch were to happen (and if they did not, revenue would fall), as well as the problem that the VCG auction is difficult to explain to advertisers (Varian and Harris, 2014).

CHAPTER 2

The Importance of Preferences in Sectoral Sorting: Direct Evidence from Lawyers in the Private and Nonprofit Sectors

2.1 Introduction

In this paper we study the market for entry-level lawyers, who make dramatically less in the nonprofit sector than in the private sector. In data from a broad survey of lawyers in 2002, two years after they graduated (the After the JD survey), the average annual difference in 2002 is $51,000. There are two broad classes of explanation for such a pay gap: differences in worker skill (ability), or differences in job characteristics (amenities) which workers value due to their preferences. We present several pieces of evidence to suggest that differences in skill are not the whole explanation, and instead we find a strong role for preferences. In our data, we have questions regarding preferences for job characteristics among lawyers in both the private and nonprofit sectors and so we are able to present some direct evidence that these preferences for job characteristics play an important role in understanding the equilibrium in this market.[1]

In particular, we show that preferences among lawyers in the nonprofit and private sectors are quite different and we attribute a large component of the private-nonprofit sector wage gap to these differences. While pay in the private sector is responsive to law school rank, pay in the nonprofit

[1] Weisbrod (1983) and Leete (2001) have previously documented that observable differences in skill do not fully explain the private-nonprofit sector pay gap. However, they do not have preference questions in their data.

sector is not. Nonprofit lawyers from higher ranked schools, who on average give up more income to join the nonprofit sector, also have stronger preferences to be in the nonprofit sector.

We pursue several complementary empirical strategies to show the importance of preferences in explaining the pay gap. First, in our primary empirical analysis, we estimate wage equations with and without preferences, under the assumption that preferences are directly correlated with private sector earnings. Second, we exploit the panel nature of our data to present within-lawyer evidence based on the subset of our sample that switches sectors. Third, we examine the subsample of lawyers who switch plans during law school. All three of these empirical strategies imply that preferences play a large role in explaining the pay gap between the private and nonprofit sectors. Our preferred estimates from our first strategy imply that entry level lawyers faced an average opportunity cost to work in the nonprofit sector of $21,000 per year, just under half of the raw wage gap, or $33,000 if they worked the longer hours typical in the private sector. We estimate that the non-hours adjusted opportunity cost ranges from $5,000 - $11,000 per year from lawyers from tier 3 and 4 law schools to $59,000 per year for lawyers from top 10 law schools (as ranked by US News and World Report).

While the theory of compensating differentials predicts that a worker should earn less in a job that has desirable amenities, researchers have had mixed empirical success in confirming this prediction. The mixed success has generated a parallel literature explaining why this prediction is not borne out in labor markets. One class of explanations focuses on issues in measuring unobserved worker productivity (e.g. Hwang et al. (1992)). Another class of explanations focuses on the presence of search frictions and firms' incentives to create amenities (e.g. Hwang et al. (1998), Lang and Majumdar (2004), Bonhomme and Jolivet (2009) and Sullivan and To (2012)).

We focus on a labor market where high quality data and institutional features make the above issues potentially less problematic. Unmeasured productivity differences are likely to be less problematic because we have reasonably detailed data on skill. Search frictions are less likely to be important in shaping market outcomes since entry-level lawyers are relatively aware of job opportunities, are relatively mobile, and the market itself is relatively well organized. While the special features of our empirical setting may limit the easy generalizability of our results to other labor markets, it does suggest that some of the empirical failures of the theory of compensating differentials are due to looking at excessively broad markets, without detailed enough data. A series of detailed case studies might paint a different portrait of the role of compensating differentials in the labor market than attempts to measure compensating differentials using less detailed datasets.[2] Though self-reported preference data raise issues of interpretation, we are reassured that the answers are representative of true preferences, because we show that they allow us to predict future behavior.

This paper makes several contributions. Our primary contribution is to show that the entry-level legal labor market presents a clear case of sectoral sorting based on underlying preferences, on which we have direct evidence. The closest paper along this dimension is Hurst and Pugsley (2011), who report survey evidence on the importance of non-pecuniary motives for entrepreneurship. However, they do not have evidence on the preferences of non-entrepreneurs and so are not able to study sorting directly.

Our second contribution is to study the nonprofit sector pay gap with explicit questions about preferences, while also attempting to account for self-selection. In previous work in this literature, Preston (1989), Leete (2001), and Ruhm and Borkoski (2003) estimate the compensating differential for nonprofit work. Preston (1989) finds some evidence of pay gaps for nonprofit work among managers after controlling for some observables but Leete (2001) finds that these gaps shrink to zero after controlling for more detailed observables. Ruhm and Borkoski (2003) argue that "competitive" pay setting is dominant, finding no strong evidence of compensating differentials, including in panel estimates of sector switchers. Weisbrod (1983) estimates the compensating differential for nonprofit legal work in particular and Leete (2001) includes case studies of particular sectors of work, including the legal sector. Both papers find that controlling for observables alone yields a large unexplained pay gap between the nonprofit and private sector that they attribute to compensating differentials.[3] Unlike Leete (2001) and Weisbrod (1983), Goddeeris (1988) attempts to explicitly account for sorting between the private and nonprofit legal sector, and finds that sectoral

[2]Indeed, Rosen's handbook chapter concludes with a call for better data: "the greatest potential for further progress rests in developing more suitable sources of data" (Rosen (1986, pg. 688)).

[3]Frank (1996) argues that the legal pay gap is due to preferences.

selection on unobserved ability can account for all of the sectoral pay differences (though the estimates are imprecise). This result is frequently cited as demonstrating the absence of compensating differentials for nonprofit work (e.g. Ruhm and Borkoski (2003); DeVaro and Brookshire (2006); Bonhomme and Jolivet (2009); and Delfgaauw and Dur (2010)).

Our data allow us to build on these papers in several important ways. A central challenge in these papers is that they do not have direct measures of preferences, yet they attribute to preferences the pay gaps that are not explained by observables, leaving open the possibility that some part of the pay gap not explained by observables is actually due to unobserved skill differences, rather than preferences. Our preference questions allow us to directly demonstrate that preference differences play an important role. There is a dispute across these papers studying the nonprofit sector about how to control for industry and occupation. In particular, Leete (2001, Table 3) argues that by adopting finer industry and occupation controls than Preston (1989), she overturns many of Preston's results. By looking within an industry and occupation, we sidestep this difficulty. Our work is thus more similar to the within-industry case studies presented in Leete (2001, Table 5) using the 1990 Census PUMS dataset, however, we use surveys of lawyers (like Weisbrod (1983), except the surveys we use are more recent, cover a broader range of lawyers, and have a much larger sample size) with targeted questions such as law school GPA, which let us better control for worker ability.

Our third contribution to the literature is to document important heterogeneity in the opportunity cost of working in the nonprofit sector by skill, which is an interesting feature of the equilibrium. We are only aware of one paper which has explored this aspect of equilibrium across the nonprofit and private sectors. Mocan and Tekin (2003) look within the child care industry and find a substantial pay premium to working in the nonprofit child care sector and that pay differences between sectors depend on skill, rather than there being a single wage gap. We estimate that lawyers from highly rated law schools face a substantial opportunity cost, whereas lawyers from lower ranked schools only give up a little to work in the nonprofit sector.

2.2 Data

We use data from two surveys of lawyers. The first, After the JD (AJD), is a nationally representative panel survey of new lawyers who first passed the bar in 2000 and has detailed information on the factors, including preferences, that led the lawyers to choose their first jobs. AJD surveyed the lawyers in two waves, first in 2002-2003 and then again in 2007-2008. The second, the University of Michigan Law School Alumni Survey (UMLS), covers one particular law school. Its key benefits are that it provides law school performance measures that are comparable across students, and it has information going back to the 1950s (based on surveys that started in 1966), so that we can follow lawyers over more of their careers and track changes in the nonprofit and private sector over time. Unlike AJD, it does not have detailed data on preferences. We use the first survey wave from AJD in our primary analysis, and use its second wave, along with the UMLS data, in additional analysis and robustness checks.

Our datasets are more detailed than the datasets typically used when estimating compensating differentials. In these datasets, all lawyers would appear to be very similar in terms of their education level (graduate) and sector, and these datasets do not have the critical questions about preferences.

See Appendix B1 for more details on the datasets and how we construct our variables.

2.2.1 Sectoral Definition

Our main grouping, "sector," classifies workers into three groups: private sector, government or nonprofit. Throughout this paper we define the nonprofit sector to exclude almost all types of government lawyers.[4] However, we do include public defenders and legal service workers in the nonprofit sector.[5] For completeness, we include the government sector in our descriptive statistics,

[4] We exclude government workers in our definition of nonprofit workers for several reasons. First, excluding government is standard in the literature: Leete (2001), Preston (1989), and Ruhm and Borkoski (2003) exclude government work. Second, the government sector has advantages relative to the private sector in terms of job security and benefits so that looking solely at wages would be misleading and we do not have substantial information about benefits from either data set. Third, as Sauer (1998) shows, the "option value" of government work is larger than for the nonprofit sector in terms of later opportunities to move into more highly paid private sector jobs.

[5] We include these workers for several reasons. First, Weisbrod (1983) and Goddeeris (1988) include them. Second, the alternative classification is to include at least some of these lawyers in government. In particular, public defenders

however, we do not discuss this sector in detail because it is not the focus of this paper.

2.2.2 Summary Statistics

Table 2.1 shows in the AJD dataset the basic empirical facts about lawyer characteristics and entry-level earnings that this paper addresses. First, lawyers in the private sector make on average just over twice as much as lawyers in the nonprofit sector, or \$51,000 more per year on average.[6] Private sector lawyers make just under twice as much per hour on average as nonprofit lawyers, because private sector lawyers work somewhat longer hours.[7] Second, while there are some differences in measures of law school quality and performance between the two sectors, it is not clear how much of the pay gap these differences explain. For example, while private sector lawyers are more likely than nonprofit lawyers to have a GPA of 3.5 or higher, nonprofit lawyers are more likely to have graduated from top 20 law schools (as ranked by US News and World Report).

Table 2.2 shows that the essential sectoral patterns found in AJD's representative sample of all lawyers are true of lawyers from the University of Michigan Law School five years after graduation as well: the private sector lawyers on average make more than twice as much as the nonprofit sector lawyers. Moreover, this pay gap is even harder to explain based on observables than in

are directly employed by the government while legal services receive some funding from government. Thus, it might be desirable to group the public defenders into government. However, the AJD does not distinguish public defenders from legal service workers. Furthermore, our discussions with various lawyers and with staff at the University of Michigan Law School Office of Career Planning suggested that public defenders and legal service lawyers do not perceive themselves to be government lawyers, and that the "type" of law student who goes into government work is very different than the "type" that goes into public defense or legal services.

[6]As a crude additional comparison, we calculate average annual earnings for 28 year olds (the average age in the 2002-2003 AJD data set, which does not vary by sector) in 2003, with a completed bachelor's degree, but no further higher education, who are working full-time, in the Current Population Survey IPUMS release. This sample makes \$46,000 per year, or slightly more than the average nonprofit lawyer in the 2003 AJD dataset. This comparison suggests that nonprofit lawyers are substantially worse off financially for going to law school, because they face 3 years of lost wages and tuition expenses in law school, and are positively selected compared to the average undergraduate student, so their wages had they not gone to law school would likely be above average.

[7]The AJD survey question asks, "How many hours did you actually work last week, even if it was atypical? (Include evenings and weekends worked.)" Contrary to our expectations, weekly hours of work and hourly wages are slightly negatively correlated for private sector workers ($\rho = -0.26$), nonprofit workers ($\rho = -0.10$) and a pooled sample of both ($\rho = -0.20$). One obvious source of bias is if highly paid lawyers report fewer hours of work than they actually typically worked, for example, if generally well paid lawyers at large law firms work more than 40 hours per week, but report 40. We examined the distribution of hours for private sector lawyers in the data and though we find bunching at 40, we also find bunching at all increments of 5 between 35 and 65, with the largest peaks at 50 with slightly smaller peaks at 40 and 60. Thus, it appears that the question effectively elicits real variation in hours, which reduces our concerns about this source of potential bias.

the AJD data, partially since the lawyers in the UMLS are more similar because they all went to the same law school. The average hours gap across sectors is smaller than in AJD. GPAs come from administrative data and are represented in standard deviations from the mean. The first year GPAs, are based on a common set of courses and are identical between the private and nonprofit sectors. The fact that the mean GPA in standard deviation units is greater than 0 among respondents suggests that higher GPA students were more likely to respond. Note that because these salary measures are taken among more experienced lawyers than in AJD 2002-2003, they are not directly comparable to those in Table 2.1.

2.2.3 Key Features of the Data

In this section we document four features of the AJD data about the entry-level legal labor market and then show that they are consistent with a simple model of sorting.

Throughout the paper, we use law school rank in 2003 by U.S. News and World Report as our primary proxy for skill.[8] Because law school admission is selective and competitive, law school rank is correlated with measures of skill.[9] In Section 2.5, we offer more detailed evidence on sectoral sorting in the setting of a single school, the University of Michigan Law School, which makes it easier to interpret data on GPA compared to looking across schools using AJD.

2.2.3.1 Lower Returns to Skill in the Nonprofit Sector

The upper panel of Figure 2.1 shows that graduates of higher ranked law schools have annual compensation distributions in the private sector that are right-shifted relative to graduates of lawyers from lower ranked schools. In particular, lawyers from top 20 law schools have salary densities that peak at around $150,000 a year, whereas lawyers from schools ranked 21 and below have an annual compensation distribution centered at just over $60,000. It is clear from this picture that in the private sector attending a better law school is associated with substantially higher average annual compensation.

[8]This is the only school ranking available in AJD, and it is grouped to protect respondent privacy.
[9]For example, LSAT scores are increasing in law school tier.

Turning to the nonprofit sector, the lower panel of Figure 2.1 shows that in the nonprofit sector the salary distributions are virtually identical regardless of law school tier, in stark contrast to the private sector. This pay differentiation by tier in the private sector but not in the nonprofit sector suggests that lawyers in the private sector are financially rewarded for skill, unlike lawyers in the nonprofit sector.

2.2.3.2 The Central Role of Preferences in Sorting

The AJD survey contains of a set of 9 questions which allow us to directly study the role of preferences in sectoral sorting.[10] These questions allow us to directly test whether sorting conforms to theory, because preferences play a central role in theories of nonprofit work and in the compensating differential literature more generally.[11]

Two of the questions about preferences are particularly important for deciding whether lawyers enter the nonprofit sector. The first asks about the importance of the "opportunity to do socially responsible work," which is generally more prevalent in nonprofit work. The second question asks about the importance of "medium to long-term earning potential."

The top panel of Figure 2.2 shows that the importance of doing socially responsible work is associated with sectoral location. In particular, among nonprofit sector lawyers, well over half reported that the "opportunity to do socially responsible work" was extremely important to them (the highest response option). In contrast, among private sector lawyers the distribution is more uniform, with less than 10% reporting that doing socially responsible work was extremely important to them.

The lower panel of Figure 2.2 shows that there is similarly strong evidence of sorting into sectors on the importance of "medium to long-term earnings potential." Lawyers in the nonprofit sector are much less likely than lawyers in the private sector to report that earnings potential is important. The bar charts are essentially mirror images of each other.

[10]The precise wording is, "Thinking about the principal types of settings in which lawyers work (e.g., government, large law firms, business), how important was each of the following factors in determining the sector in which you began your professional career?"

[11]See Preston (1989), Frank (1996), and Delfgaauw and Dur (2010) as examples of models of nonprofit pay where preferences play a central role. Rosen (1986) is the classic reference on compensating differentials where preferences play a key role.

Table 2.3 shows the results of a probit model on selection into the private and nonprofit sectors to explore the roles of the skill, demographic and preference variables together. Column (1) does not include the preference variables, then we add social responsibility preference to column (2), and then further add earning potential preferences to column (3). Overall, this model suggests that preferences play a much stronger role in sector choice than other possible factors.

In the first column, the point estimates suggest that private sector lawyers are somewhat negatively selected on law school tier (e.g. a lawyer from a top 10 school is 3 percentage points less likely to join the private sector compared to a lawyer from a tier 4 school). Though none of the estimates are statistically significantly different from 0 individually, a χ^2 test of their joint significance rejects the null that they are all equal to 0 with a test statistic of 10.05 (p=0.040). In contrast, the point estimates suggest that private sector lawyers are somewhat positively selected on GPA. Lawyers with GPAs above 3 are about 6-8 percentage points more likely to join the private sector than lawyers with GPAs below 3 (although we see from Table 2.1 that the vast majority of lawyers have GPAs above 3. The χ^2 test of the joint significance of the GPA variables rejects the null with a test statistic of 15.03 (p=0.010). Additionally, men are 7.2 percentage points more likely to join the private sector, and this estimate is highly significant. Overall, these factors alone are not tremendously predictive of sector choice. The pseudo-R^2 of the model is 0.110.

In contrast, the social responsibility and earnings potential preference variables are much more predictive. Adding them sequentially to columns (2) and (3) raises the pseudo-R^2 to 0.307 and 0.373, respectively. χ^2 tests of the significance of these sets of variables are highly significant, with $p < 0.0001$. In column (2), we see that a lawyer with the strongest stated preferences to do socially responsible work is 37 percentage points more likely to join the nonprofit sector, compared to a lawyer with the lowest stated preferences. Similarly, in column (3), we see that a lawyers with the strongest preferences for earning potential is 18 percentage points more likely to work in the private sector. Adding the earnings potential preference variable does not significantly change the estimated effects of the social responsibility variable. Adding both preference variables does not substantially change any of the estimates from column (1), with the small exception of lowering the marginal probability differential between men and women to 3.1 percentage points.

While the other preferences questions are not our main focus, there are some differences in responses by sector. Table 2.4 shows mean responses to all of the preference questions by sector. The ability to pay off debt is most important to private sector workers, while the importance of loan repayment programs is highest for nonprofit workers. These two patterns are unsurprising. Private sector lawyers value prestige and future career mobility more than nonprofit lawyers. If anything, these two factors are an additional benefit of having a private sector job. Finally, nonprofit lawyers (and government lawyers) value work-life balance more than private sector lawyers, which makes sense considering that private sector lawyers spend more time at work each week.[12]

2.2.3.3 Higher Skill Lawyers in the Nonprofit Sector Have Stronger Preferences

The right hand side panel of Figure 2.3 shows that the average importance of doing socially responsible work among lawyers in the nonprofit sector is increasing in the quality of the law school. This increase is relatively small, however, which is not surprising given that Figure 2.2 shows that there is little variation in the response to this question among lawyers in the nonprofit sector. The other preference variable, importance of earnings potential, displays more variation within the nonprofit sector both overall and by school tier. It shows that lawyers in the nonprofit sector from higher-ranked schools value earning potential substantially less than lawyers from lower ranked schools: the average importance falls from a 4 to a 2 as we move from schools ranked 101 and up to top 10 schools. The preferences over earnings potential are strongly, negatively correlated with law school rank among nonprofit sector lawyers.

A potential concern is that these patterns simply reflect differences in preferences common across all lawyers from each tier of law school. The left hand side panel shows that this is not the case: these preferences do not vary substantially across tiers in the private sector.

[12]One potential downside of the preferences questions is that they were asked in 2002-2003, at least two years after the lawyers first decided where to work. This lag means that the responses might not reflect the lawyers' ex-ante preferences when they were making a career decision, but instead might reflect an ex-post justification for their early career sector choices. Thus, the reported preferences might be uninformative about the basis of past decisions. We show in Section 2.5 that the responses to preference questions in 2002 are correlated with movements between sectors after this survey was conducted, so they are at least informative about future decisions, suggesting that they are correlated with respondents' true underlying preferences. Furthermore, respondents do not merely give the "polite" answers of caring a lot about social responsibility and little about earning potential. Thus, we have at least some evidence that respondents' responses reflect their true career preferences.

2.2.3.4 Distribution of Preferences are Independent of Law School Rank

Table 2.5 shows that the distribution of preferences in the AJD survey is roughly independent of the strongest correlate we have of skill, law school rank. It shows the average response to the two questions that we showed are important in determining sectoral location: importance of the opportunity to do socially responsible work and medium to long-term earnings potential in your choice of job sector. The table shows that the mean and standard deviation of the responses to both questions are roughly constant across law school tiers.

2.3 A Conceptual Model of Sorting

The key features of the data we document are consistent with sorting on the basis of preferences under one additional assumption: returns to skill are lower in the nonprofit sector than in the private sector. The intuition is that because high-skill lawyers give up more pay to locate in the nonprofit sector, those lawyers who do so will have stronger preferences to be there. To make this intuition precise, this section develops a sorting model of a labor market with preference heterogeneity. The model also highlights the fact that compensating differentials are heterogeneous in the presence of such differentiated opportunity cost. We discuss the notion of the opportunity cost that we estimate in the remainder of the paper as the difference between realized wages in one sector and counterfactual wages in another.

2.3.1 Environment

Our model is a Roy (1951) model of sorting between the nonprofit and private sector on the basis of comparative advantage in producing utility. Lawyers have a level of skill and a preference for working in the nonprofit sector. The private sector pays more for skill than the nonprofit sector, while the nonprofit sector provides more non-wage benefits from the type of work lawyers do in it.

Formally, each lawyer is endowed with skill denoted by X, which is a random variable, that we assume is perfectly transferable across sectors. Most generally, X may either be a scalar or a vector of observed and unobserved traits. Each lawyer also derives utility ϕ from working in the nonprofit sector, which is also a random variable. Consistent with our facts, let X and ϕ be independent.

Lawyers choose to work in either the private or nonprofit sector, which yield utility values u_1 and u_0, respectively. In the private sector, a lawyer would gain utility from their salary (w_1), which depends on their skills, returns to skills (β_1) and a random draw of earnings, ϵ^1 so that the potential utility in the private sector is:

$$u_1 = w_1 = X\beta_1 + \epsilon^1,$$

where $\mathbb{E}[\epsilon^1] = \mathbb{E}[\epsilon^1|X] = 0$. In the nonprofit sector, a lawyer has different returns to skills (β_0) and has a modified intercept (δ), thus earning a salary of $w_0 = X\beta_0 + \delta$ and they also receive additional non-salary-based utility, ϕ, as well as a random utility draw ϵ^0 so that total utility is:

$$u_0 = X\beta_0 + \delta + \phi + \epsilon^0.$$

2.3.2 Equilibrium: Sorting and Selection

A lawyer chooses a sector to maximize utility. In particular, a lawyer locates in the nonprofit sector if $u_0 > u_1$ or equivalently:

$$\phi > X(\beta_1 - \beta_0) - \delta + (\epsilon^1 - \epsilon^0) \equiv \bar{\phi}. \tag{2.1}$$

If X is a vector, the model allows for comparative advantage, since some characteristics might be more highly rewarded in one sector compared to another. Going forward though, since pay is relatively flat in the nonprofit sector, we think of X conceptually as a scalar composite of skill variables, ruling out comparative advantage.

Figure 2.4 depicts the resulting sorting of lawyers under two further simplifying assumptions: $\beta_1 > \beta_0$, that is, returns to skill are higher in the private sector (as documented in Section 2.2.3.1) and $\epsilon^1 = \epsilon^0$ (an assumption used here only to simplify the figure). Lawyers endowed with more skill relative to their preferences sort into the private sector, while lawyers with stronger preferences relative to their skill sort into the nonprofit sector.

Because preferences are not proportional to skill, the key feature of the equilibrium is that the cutoff preference intensity to locate in the nonprofit sector is increasing in skill. This feature has two implications. First, lawyers with more skill who are in the nonprofit sector have stronger preferences to be there. Second, combining the rising preference cutoff with the fact that the distribution of preferences are roughly independent of skill, there are fewer high skill lawyers in the nonprofit sector and so the nonprofit sector is negatively selected.

2.4 Main Results

2.4.1 Empirical Model of Private Sector Wages

We develop a simple empirical model to predict lawyer wages in the private sector as a function of ability, using law school tier and GPA as our primary predictors, which we then use to predict counterfactual private sector wages for the nonprofit lawyers. We model log hourly wages in the private sector as:

$$y_i = X_i\beta + \epsilon_i \tag{2.2}$$

where y_i is log hourly wages in the private sector, X_i are ability variables and ϵ_i is an error term.

If sector choice is uncorrelated with ϵ then Equation 2.2 not only models wages for private sector lawyers, but also models counterfactual wages for lawyers who choose to work in the nonprofit sector, if they instead worked in the private sector. This counterfactual represents the wages nonprofit lawyers would earn in the private sector if they earned the same wages as an equivalently skilled private sector lawyer. We are unable to test this assumption directly and we test sensitivity to this assumption in Section 2.4.4.

2.4.2 Wage Equation Estimates By Sector

Using our sample of private sector lawyers from AJD, we estimate Equation 2.2 via OLS, and show the results in Table 2.6 in column 1. Table 2.6 shows that skill variables are strongly associated with wages. Each higher level of law school rank and GPA is generally associated with an increase in salary: a 55 log points increase for a top 10 school relative to tier 4 law school, and a 50 log point increase for a GPA of 4-3.75 relative to 2.49-2.25. Being on a general or special law review and being male are associated with higher wages, although the magnitudes of these effects are smaller than those of law school rank and GPA.

Table 2.7 shows the analogous results within the nonprofit sector. The patterns are very different, although the smaller sample size of this sector makes most of our estimates less precise. There is no longer strong evidence of increasing wages in law school tier and GPA.

Table 2.8 shows similar cross-sectional estimates using the UMLS data for both sectors, 5 and 15 years after law school. Due to the smaller sample size, we pool over multiple class years, include a linear graduation year trend, and use real earnings. The estimates from the UMLS fit a key pattern we found using the AJD data: pay increases with GPA in the private sector, but not significantly in the nonprofit sector. We estimate that 5 and 15 years after law school, a University of Michigan lawyer in the private sector earns hourly wages 11 and 17 log points higher per unit increase in the standard deviation of their first year GPA and these estimates are highly significant and fairly precise. In contrast, the GPA point estimates in the nonprofit sector are about half as large, are not significant since they are far less precise.

2.4.3 Estimates of Opportunity Cost of Working in the Nonprofit Sector

We calculate a counterfactual expected private sector wage for each nonprofit worker ($\mathbb{E}[y_i|S_i = 0]$) using Equation 2.2. Using this counterfactual, we calculate an "individual opportunity cost" (OC_i) as the amount in log hourly wages we predict a nonprofit lawyer gives up to work in the nonprofit sector instead of the private sector as:

$$OC_i = \mathbb{E}[y_i | S_i = 0] - y_i^{\text{nonprofit}} \tag{2.3}$$

where $y_i^{\text{nonprofit}}$ are actual wages for nonprofit lawyers. We thus can calculate mean opportunity costs based on the estimated empirical distribution of individual opportunity costs.

In Table 2.9, row 1 reports a mean estimated average opportunity costs of 45 log points based on the estimates from column (1) in Table 2.6 using the AJD data. The mean raw log hourly wage difference in the AJD data between the two sectors is 55 points. The 10 log point difference between our estimated mean opportunity cost and the mean sectoral wage difference suggests that nonprofit lawyers would earn less in the private sector than actual private sector lawyers. That is, nonprofit lawyers are negatively selected on skill (following our model, Section 2.3.2), but not nearly as much as the raw wage gaps predicts. Although this log points difference is small relative to the raw wage difference, we reject the hypothesis that it is 0 and we also reject the hypothesis that it is 55, the raw difference between the two sectors.

We also present estimates of opportunity costs using the UMLS data based on the estimates in Table 2.8. In Table 2.9, rows 2 and 3, we estimate the opportunity cost of working in the private sector 5 and 15 years after law school to be 64 and 78 log points, respectively. These estimates are fairly precise and highly significant. Like the AJD estimates, these estimates suggest that lawyers give up substantial potential earnings to work in the nonprofit sector.

We next translate these hourly wage estimates using the AJD data from logs into levels, and then into annual wage estimates. Average hourly wages for nonprofit and private sector lawyers are \$21 and \$40 respectively. Average annual wages are \$44,000 and \$95,000, respectively. Using the logs-to-levels retransformation in Duan (1983), we estimate counterfactual average hourly wages of \$31 for nonprofit lawyers if they instead worked in the private sector. Nonprofit lawyers work fewer hours per week on average than private sector lawyers, 40.5 hours compared to 48.1 hours. At our estimated counterfactual hourly wage, in the private sector, nonprofit lawyers would make an average of \$75,000 per year working nonprofit hours, or \$90,000 per year working private sector hours.

In Table 2.10, we show estimated average opportunity costs for nonprofit lawyers by law school rank using the AJD data. These estimates support the theory that lawyers from stronger law schools give up more pay to join the nonprofit sector. We estimate that lawyers from top 10 law school give up 106 log points on average in hourly wages to join the nonprofit sector, which is significantly different from the average opportunity cost of 45 log points. Our estimate of lawyers' average opportunity cost from tier 3 and 4 schools is much smaller, 16 and 20 log points respectively, and both are significantly different from both 0 and the average opportunity cost. Thus our results suggest that only lawyers from highly ranked schools give up substantial earnings to work in the nonprofit sector.

2.4.4 Robustness of Estimated Opportunity Cost to Selection on Unobservables

We are concerned that if nonprofit lawyers are negatively selected on unobservable factors which influence potential private sector wages (as in the model of Goddeeris (1988)), then we would be overestimating their opportunity costs. We do not believe that our data include a credible instrument for sector selection, so instead we check the sensitivity of our estimates to selection on unobservables. As a robustness check, we adopt the methods of Altonji et al. (2005) to explore potential bias from sectoral selection on unobservables. We assume that there is potentially correlation ρ between the error terms in the selection and wage equations, but, lacking a suitable instrument, the correlation is not well identified. We estimate the Heckman (1979) model using MLE, assuming bivariate normal error terms and that the preference variables belong in both the wage and selection equations, but we constrain the correlation between the equations to a range of fixed values.[13]

We begin with a standard use of the law of total expectation to derive expected counterfactual wages for nonprofit workers:

$$0 = \mathbb{E}[\epsilon] = \mathbb{E}[\epsilon|X] = P\mathbb{E}[\epsilon|P, S = 1] + (1 - P)\mathbb{E}[\epsilon|P, S = 0] \tag{2.4}$$

where P is the propensity to work in the private sector. We rearrange as:

$$\mathbb{E}[\epsilon|P, S = 0] = -\frac{P}{1 - P}\mathbb{E}[\epsilon|P, S = 1] \tag{2.5}$$

We thus construct counterfactual expected wages for nonprofit workers as though they worked in the private sector as:

$$\mathbb{E}[y_i] = X_i\hat{\beta} - \frac{\hat{P}}{1 - \hat{P}}\mathbb{E}[\epsilon|\hat{P}, S = 1] \tag{2.6}$$

using $\hat{P}$ and $\hat{\beta}$ as estimated by the constrained Heckman model.

Table 2.11 shows the resulting mean opportunity costs for a range of ρ values between -0.5 and 0.5. As expected, estimated opportunity cost is decreasing in the correlation between the unobservables in the selection and wage equations. A positive correlation between these terms means that lawyers who are more likely to work in the private sector are positively selected into this sector (e.g. they have favorable unobservables), which means that that observably equivalent

[13] We also considered an alternative Heckman (1979) style model which allows sector selection based on unobservables, where we assume that the preference variables meet the necessary conditions to be excluded variables to correct for sectoral selection. We estimated the model using the Heckman two-step method, as well as a semi-parametric method which builds off of Chamberlain (1986); Heckman (1990); Cosslett (1991); Vella (1998); Newey (2009); French and Taber (2011). This estimation strategy is similar to the one used by Goddeeris (1988). We are not convinced that preference only influence wages via their influence on sector selection and the estimates from this model suggested that nonprofit workers would make implausibly higher wages in the private sector than private sector workers. In our data, we do not believe there are credible instruments for sector selection. We thus do not present these results.

nonprofit lawyers would have relatively lower wages in the private sector. Assuming no correlation between the equations (ρ=0) is equivalent to the estimate in Table 2.9, row 1. In the AJD column, the point estimate of the average opportunity cost remains positive for all values of ρ in our range, and remains significantly for $\rho < 0.5$. The UMLS column shows very similar results. Thus, there is evidence that selection on unobservables would need to be quite substantial to reverse our result that nonprofit lawyers on average would earn more if they instead worked in the private sector.

2.4.5 The Role of Preferences Within the Private Sector

Jobs vary within the private sector not only by pay, but also area of law, types of clients and other work conditions. We consider a versions of the model that includes our preference variables (preferences for socially responsible work and earning potential) as predictors of wages within the private sector:

$$y_i = X_i\beta + Z_i\alpha + \epsilon_i \tag{2.7}$$

where Z_i are preference variables. When we include these variables, we interpret them as representing choice of job characteristics within the private sector.

We consider estimates of Equation 2.7 in Table 2.6, including socially responsible preferences (column 2), and both preference variables (column 3). Preferences do in fact influence wages. When we include social responsibility preferences (column 2), the estimates on the ability terms do not change significantly, and we estimate that private sector lawyers with strong social responsibility preferences earn less. In particular, we estimate that the lawyers who care the most about social responsibility (those with a 7) earn 21 log points less than lawyers who case the least (the omitted group, those with a 1). When we next add earning potential preferences to our model (column 3), the estimates of the other coefficients again do not change significantly. We estimate that lawyers with strong earnings potential preferences earn more: lawyers with a 7 earn 29 log points more than those with a 1 (the omitted group). Both sets of social responsibility and earnings potential preference are significant in tests that all of their coefficients are equal to zero, with F-statistics of 3.45 (p=0.0022) and 4.10 (p=0.0003) respectively. These estimates suggest that preferences have a strong influence on earnings for lawyers in the private sector.

2.5 Fixed Effect Analysis of Selection

2.5.1 Analysis of Sector Switchers

In this section we pursue a strategy to control for unobservables to augment our primary analysis: we exploit the panel structure of both of our datasets to look at lawyers who switch sectors over time, which allows us to estimate within-lawyer sectoral pay gaps and thus opportunity costs (in log wages). We use standard panel data methods. This analysis supports our main results, and though conceptually we gain power though paired observations of the same lawyers, we have small sample sizes of sector switchers, which reduce power and thus limits our analysis somewhat.

If a random sample of lawyers switched sectors, then a fixed-effects model of within-lawyer wage differences, including years after graduation fixed effects (at either 5 or 15 years after the initial survey year) would provide an unbiased estimate of the average opportunity cost of working in the nonprofit sector as a more experienced lawyer. However, lawyers could switch non-randomly as a function of their potential wages in both sectors, potentially biasing our estimates, although since wages do not vary that much in the nonprofit sector, we are primarily concerned about variation in actual or potential wages in the private sector. Our estimates would not directly be biased if lawyers switch non-randomly as a function of preferences, if preferences are unrelated to their potential wages in both sectors.

The sector switchers also allow us to examine two other limitations of our previous analysis. First, in Section 2.2.3.2 we were concerned that responses to the preferences questions in the AJD survey might have reflected ex-post feelings or justification for career choices. Here we show evidence that the preference responses are valid since the lawyers who switch sectors between the two waves of AJD reported systematically different preferences in the first wave than those who do not. The second limitation of our previous analysis is that we do not observe the actual law school or first-year GPA of AJD respondents. With sector switchers we can use our data from UMLS and exploit the fact that the first-year GPA data is from the same set of classes at the same law school to see how selection works on the basis of GPA.

2.5.1.1 After the JD

In After the JD, between 2002-2003 and 2007-2008, 17 lawyers switched from the nonprofit to the private sector, and 29 switched from private to nonprofit.[14] Table 2.12 shows the summary statistics on the subsample of sector switchers.

The table makes three points. First, switchers in either direction do not appear to be substantially selected on pre-switching sectoral earnings. In 2002, prior to switching, the lawyers who eventually switched from the nonprofit to the private sector made 10 log points more in hourly wages than the lawyers who did not. Prior to switching, the lawyers who switched from the private sector to the nonprofit sector made 11 log points less.

Second, the movements in pay among the switchers reveal important sectoral pay gaps, even within an individual. In 2007, the nonprofit sector lawyers who switched to the private sector made 67 log points more than the lawyers who stayed in the nonprofit sector. Similarly, the private sector lawyers who moved to the nonprofit sector saw their wages made 45 log points less than the lawyers who stayed. These post-switching differences are far larger than the small pre-switching differences between these groups.

More formally demonstrating the difference-in-difference approach, Table 2.13 shows two fixed-effects regressions, including individual and wave fixed effects. The independent variable of interest is the sector switching indicator in wave 2. Column (1), estimates that switchers from the nonprofit to the private sector earn 57 log points more in hourly wages in 2007. Similarly, Column (2) estimates that switchers from the private to the nonprofit sector give up 34 log points in hourly wages in 2007. Both estimates are highly significant.

Third, reported preferences in 2002 are correlated with movements between then and 2007. The lawyers who left the nonprofit sector reported weaker preferences to do socially responsible

[14] These are the observations for which there is wage data in both waves which also meet the other restrictions in Appendix B1.

work and stronger preferences for earnings potential than the lawyers who stayed. Similarly, the lawyers who left the private sector reported stronger preferences to do socially responsible work and valued earnings potential less than the lawyers who stayed in the private sector. These findings suggest that the preference questions are meaningfully correlated with actual behavior and future actions. They thus provide some evidence that the responses to the preference questions (in 2002) reflect true sentiments about preferences at the time lawyers made their initial sector choice (in 2000), rather than solely being ex-post justification.

In combination, the switchers in both directions suggest the presence of large compensating differentials for sectoral choice among lawyers. This finding is in contrast to the results in Ruhm and Borkoski (2003) who find in the CPS that switchers in both directions earn small wage increases, however, our results are not directly comparable. Though they analyze some occupations, they do not specifically analyze lawyers.

2.5.1.2 University of Michigan Law School

The UMLS data also has sector switchers and we present a similar analysis for sector switchers after 5 years. Since UMLS is a longer panel, we also present evidence from 15-year sector switchers. Because of the smaller sample sizes of UMLS, we created pooled samples of 10 graduation years (1991-2000 for the 5-year switchers and 1981-1990 for the 15-year switchers) and report real (2002) wages, adjusted by the CPI. Unfortunately, only annual, not hourly earnings (nor hours of work) are available in the first year, so we do all comparisons in log annual wages. However, we feel this is a minor concern because differences in weekly hours of work between the two sectors are small in the UMLS data (Table 2.2). The results are substantively similar to those from AJD.

The UMLS data allows us to examine two issues that the AJD data does not. First, we can examine longer-run salary differences between the sectors because we have a longer panel in UMLS. Second, we can examine sorting on the basis of law school performance using within-school measures of GPA.

Table 2.14 makes four points about the salaries. First, as in the AJD data, the switchers from nonprofit to private are not systematically selected on potential earnings. If anything, the switchers are negatively selected based on initial salary. For both sectors, those who eventually switched made less initially. Second, switching from the nonprofit to the private sector yields substantial wage increases relative to the stayers, and the raw differences are especially large 15 years out, at 47 log points ($92,000) a year. Third, salary decreased for lawyers who switched within five years from the private sector to the nonprofit sector, while the analogous fifteen year switchers experienced a relatively small increase in earnings. However, both the five and fifteen year switchers away from the private sector make far less in the nonprofit sector than the lawyers who stay in the private sector. Finally, as with the AJD data, the switchers make clear that the way to maximize long-term earnings is to start in the private sector and stay there. Fifteen years into their careers, lawyers who started and stayed in the private sector make over 73 log points more than those who start in the nonprofit sector and switch to the private sector.

Table 2.15 shows the same fixed-effects regressions using the UMLS data as Table 2.13 shows with the AJD data, yielding very similar results. Column (1) estimates that after 5 years, switchers out of the nonprofit sector earn 28 log points more, although this estimate is not significant at conventional levels (p=0.073), it is in the same direction as the comparable estimate from AJD. Column (2) estimates that after 5 years, switchers out of the private sector earn 54 log points less. This estimate is significant, and larger than the comparable one using the AJD data. Columns (3) and (4) show the same set of estimates, using 15 years after law school as the second period, and the estimated effects of switcher are larger after 15 years (69 and -75 log points, respectively) and significant.

Table 2.14 makes two points about GPA. First, as we expect from the model, there is some negative selection into the nonprofit sector on the basis of both first year and overall GPAs. Second, however, this negative selection does not confound our analysis of switchers. The lawyers who initially locate in the nonprofit sector and switch to the private sector have substantially lower GPAs than those who stay, both 5 and 15 years out, while the evidence from GPAs on those who switch from the private sector to the nonprofit sector is more mixed, because 5 year switchers have slightly lower GPAs, while 15 year switchers have slightly higher GPAs. As with the AJD switchers, the evidence from the UMLS switchers suggests that there is a large opportunity cost to working in the nonprofit sector.

2.5.2 Change of Sector Plans During Law School

The primary empirical concern with interpreting sectoral pay gaps as primarily reflecting preference, rather than ability differences, is that sorting on unobserved ability might be the main driver of sectoral pay differences. In the previous subsection, we explored pay gaps within-lawyers who switched sectors once already in the labor force as a way of controlling for unobserved ability. Here we explore two additional aspects of selection: selection on GPA within a particular law school, and also selection on GPA of lawyers who change their planned long-term sectoral location while in law school. The goal in both cases is to see the extent to which lawyers who eventually end up in, or switch towards, the nonprofit sector are negatively selected.

The UMLS survey asks lawyers about what their long-term career plans were when they entered and exited law school. We also know their law school GPAs at the end of their first year and their cumulative GPA. Therefore, we can see if law students systematically switch into either the private sector or nonprofit sector based on their performance during law school.

This analysis is similar to work by Stinebrickner and Stinebrickner (2011) on college major choice. They show that students drop out of math and science majors as they get negative feedback about their performance and update their beliefs about their own abilities, suggesting that math and science get a positive selection of graduates.

The plans switchers suggest that, if anything, there is a positive selection of lawyers into the nonprofit sector. Table 2.16 shows the GPAs (in standard deviation units) of lawyers by original sector plans and the sector they ended up in. The lawyers who keep their plans of staying in the nonprofit sector did better in law school (with first year and final GPAs of 0.30 and 0.42 standard deviations above the mean, respectively) than those lawyers who started out planning on going into the nonprofit sector and switched to the private sector (who had first year and final GPAs of -0.04 and 0.02 standard deviations from the mean, respectively).[15] Similarly, those students who

[15] Since nearly all UMLS students take the same first-year classes, we focus on first year GPAs to avoid issues of sorting into easier and harder classes in years 2 and 3. As Table 2.16 shows though, final GPAs are very similar to first-year GPAs.

entered planning on working in the private sector but ended up locating in the nonprofit sector did better in law school (with an average first year GPA of 0.21 standard deviations above the mean) than those students who stayed in the profit sector (with an average first year GPA of 0.16 standard deviations above the mean). Regardless of which sector the students initially planned on going into, the nonprofit sector appears to get a positive selection of students.[16] We thus find no evidence consistent with endogenous, positive selection on revealed ability into the private sector based on UMLS students' changing career plans during law school.

2.6 Alternative Potential Explanations for the Sectoral Pay Gap

While we have documented substantial pay gaps between the nonprofit and private sectors and interpreted these gaps as at least partially representing opportunity costs due to preference differences, we consider several alternative potential explanations. None of them refute the notion that nonprofit lawyers give up substantial wages to be in the nonprofit sector.

2.6.1 Lifetime Earnings

The entry-level wages that are the primary focus of our analysis, while convenient to use given the available data, are not the best earnings metric if we think that lawyers value and optimize their appropriately discounted lifetime earnings. If low entry wages for nonprofit lawyers lead to higher wages later on than other career paths, then our analysis would overstate the opportunity cost when viewed from a lifetime earnings perspective. However, our evidence suggests that the entry level sectoral pay differences understate the relative lifetime ones.

We make two points. First, the salary changes of switchers from the nonprofit to the private sector show that they make less than those who had started in the private sector, suggesting that starting elsewhere and switching to the private sector is not a good strategy for salary maximization. The evidence on sector switchers from AJD and UMLS in Tables 7 and 2.14 suggests that the way to maximize later-career earnings is to start out in the private sector and stay there. The average

[16] As with the UMLS graduates from 1996-2001 in Table 2.2, the 1982-1991 graduates who responded to the survey have slightly higher than average GPA's.

later-career salary in both data sources from only working in the private sector is substantially higher than salaries for other career paths. The lawyers who start out in the nonprofit sector and switch to the private sector do experience substantial pay increases, but they do not close the gap with those who start out in the private sector.

Second, the lifetime income differences conditional on initial sectoral location are substantial. In the UMLS dataset we have observations on lawyers 1, 4, 14 and 24 years after law school, for the sample who graduated between 1976-1979, and we use these observations to estimate lifetime income.[17] The implied lifetime income gap (for the first 24 years of legal work) based on initial sectoral location is large: the nonprofit starters have a lifetime income of about $1,000,000 and the private sector starters have a lifetime income of about $2,400,000. This comparison demonstrates a substantial lifetime earnings gap between lawyers based on their original sectoral choice and likely underestimates the true lifetime earnings gap by not observing wages late in these lawyers' careers. Furthermore, because the private-nonprofit sectoral wage gap has grown substantially since the 1970s and especially during the 2000s, we expect the lifetime earning gap for more recent law graduates to be even larger.

2.6.2 Debt Forgiveness

There are state, federal, and law school specific loan forgiveness and repayment programs for lawyers, which have varied over time. These programs provide assistance in various ways, including: ongoing debt assistance payments, reduced required monthly loan payments, and complete debt forgiveness after a certain number of years. Depending on the program, lawyers qualify for assistance by having a low income level and/or a job in the nonprofit sector (or in some cases, the government sector). Typically, for assistance in the form of ongoing debt payments, the size of these assistance payments decreases with a lawyer's income.

These programs somewhat complicate our comparison of salaries between the nonprofit and private sectors because typically a lawyer would receive more assistance in the nonprofit than the

[17] We convert each of these income observations to 2002 dollars and then impute salaries for the missing years by geometrically interpolating the two nearest observations. We then discount this income stream back to the year of graduation using a discount rate of 0.95.

private sector. This difference, if ignored, would bias us towards overestimating the opportunity cost of working in the nonprofit sector, but is only a major concern if it is actually large. Ideally, we would estimate the value of loan assistance programs on an individual level to the lawyers in our AJD sample, and their value if the lawyers instead chose the other sector, and we would compare this estimated difference in assistance between the sectors to our estimated opportunity costs. Unfortunately we are unable to directly estimate loan assistance values at the individual level, due to a combination of factors: the overall complexity of these programs, their changes over time and the fact that we do not have available the specific law school a lawyer attended in the AJD data. However, the evidence we do have suggests that loan repayment programs are unlikely to have a major impact on our estimates.

We make three points. First, the magnitude of debt, which generally is an upper bound on the value of forgiveness programs, is dwarfed by the lifetime income differences between the sectors, so debt forgiveness programs cannot go far in explaining the sectoral pay gaps. Tables 2.1, 2.2 and 2.17 show the average level of education-related debt (including undergraduate loans) at the time of law school graduation by sector and law school rank, and it does not vary much by either factor. In the previous section we documented that the lifetime pay differences are on the order of $1,000,000 dollars (likely an underestimate) in present discounted value terms, while the magnitude of debt is on the order of $60,000. So, the magnitude of debt is inconsequential relative to the income differences and thus debt and assistance programs do not seem like they could plausibly be playing a central role in sectoral sorting.[18]

Second, the nonprofit lawyers generally do not report that debt assistance programs had a major impact on their sector choice. One of the survey questions about sectoral choice introduced in Section 2.2.3.2 asks about the importance of the "availability of loan repayment assistance or

[18]Two recent papers present evidence of a behavioral response to student debt that does not fit within standard models. Their findings could argue against this point. Field (2009) documents that the mere framing of loan forgiveness programs for NYU law students—whether as a loan forgiveness program conditional on sectoral location, or as a tuition remission conditional on sectoral location—has a quantitatively significant impact on their probability of going into the nonprofit sector. Rothstein and Rouse (2011) show that debt causes undergraduates in a highly selective university to avoid public interest jobs, using evidence from a policy shift of replacing all student loans with grants. However, there is also behavioral evidence supporting the notion that debt assistance programs might not play a role in sector choice. Dynarski and Scott-Clayton (2006) finds that the complexity of undergraduate financial aid programs hinders their intended use of inducing marginal students to go to college. Likewise, we suspect that the complexity of loan assistance programs for lawyers reduces their impact on sectoral choice.

loan forgiveness programs". As shown in Table 2.4, private sector lawyers rate these programs of low importance to them, with an average of 1.89 (where 1 is "Not Important at All" and 7 is "Extremely Important."). Perhaps surprisingly, nonprofit sector lawyers give these programs a 3.65 on average, indicating that they are not especially important to their choice of sector either. Nonprofit lawyers from different law school tiers have similar average responses, except for lawyers from top 10 schools, who have an average response of 6.3, indicating that they found these program important.[19] However, the annual sectoral pay gap is so large for lawyers from top 10 schools, at about $100,000, that our argument that the lifetime earnings gap dominates the size of loans and any possible assistance especially applies to these lawyers.

Finally, sectoral sorting over time has been insensitive to an increase in private sector annual wages roughly equal to the entire average debt burden, among UMLS students. Over the 25 years in our data, UMLS entry-level private sector wages have roughly doubled (in 2002 dollars), from $60,000 to $120,000, while the entry-level nonprofit wages have remained relatively flat. This change implies that the present discounted value of lifetime income lost to entering the nonprofit sector has risen dramatically over this period, and by an amount far larger than that forgiven by debt forgiveness programs. If debt forgiveness programs were playing a key role in inducing sorting, then we would expect the increase in the pay gap to reduce the number of lawyers entering the nonprofit sector, or to change the nature of selection. However, as shown in Figure 2.5 the share of UMLS students entering the nonprofit sector has not shown a systematic relationship to these pay gaps (and has been relatively flat over time). The GPA of UMLS students entering the nonprofit sector has varied somewhat over this period, but if anything, it has increased in recent years as the pay gap has increased.

2.6.3 Differences in Family Background

An alternative concern about our preference variables is if they are correlated with an important omitted determinant of sector choice. In particular, if preferences are correlated with family background, and particularly financial educational support, then family background may play a role

[19] Higher-ranked law schools tend to have more generous loan assistance programs.

similar to loan forgiveness programs, except perhaps on a larger scale.

However, there are not substantial differences in family financial support or background between the sectors. Table B3 shows how lawyers from AJD in different sectors paid for law school. Though there are slight differences between the sectors and the response rate is low to these questions, overall the share of financial support from difference sources does not vary much by sector. Federal loans are by far the largest source of support and are a slightly larger source for nonprofit lawyers than private lawyers, who received slightly more support from parents and from employment (possibly through summer internships). Table B4 shows that there are only very slight differences in education levels of lawyers' parents between the sectors.

2.7 Conclusion

In this paper we have explored the role of preferences in determining labor market outcomes. In particular, we have demonstrated that preferences strongly influence the sorting of lawyers between the nonprofit and private sectors. Private sector lawyers have higher salaries than nonprofit lawyers, yet, they are similarly skilled. We show that these preference differences explain a share of the pay gap. The AJD data set we used contains excellent survey questions for studying the role of preferences in sorting, and its use is new to the literature on preferences and sector choice.

Our most striking finding is that this pay gap is largest for higher-skilled lawyers because the nonprofit sector appears not to financially reward skill, while the private sector does. We presented evidence that there is correlation among lawyers in the nonprofit sector between skill and preferences for being in the nonprofit sector. Together, these results imply that the opportunity cost of joining the nonprofit sector for lawyers is heterogeneous in skill. This result may not be unique to our setting and heterogeneous compensating differentials may be common for other non-wage amenities and in other parts of the labor market.

Table 2.1: Lawyer summary statistics by sector, After the JD

	Private	Gov	Nonprofit	Overall
Annual Wage	$94,745	$52,400	$44,056	$84,977
	(40,992)	(15,295)	(14,343)	(41,263)
Weekly Hours	48.1	41.8	40.5	46.6
	(9.1)	(6.0)	(6.4)	(9.0)
Hourly Wage	$40.5	$25.3	$21.8	$37.0
	(22.2)	(7.0)	(6.3)	(21.0)
Debt	$64,728	$62,576	$66,974	$64,501
	(40,305)	(38,319)	(41,097)	(40,008)
Law School Ranking, % from tier:				
Top 10	10.0	4.2	10.2	9.0
Top 11-20	12.8	6.5	18.8	12.1
Top 21-100	48.0	56.3	36.7	48.7
Tier 3	16.9	16.9	19.5	17.1
Tier 4	12.4	16.1	14.8	13.1
Law School Performance, % with GPA:				
3.75 - 4.00	8.3	3.9	5.5	7.4
3.50 - 3.74	20.2	14.6	13.3	18.9
3.25 - 3.49	27.4	22.7	26.6	26.6
3.00 - 3.24	28.1	28.6	24.2	28.0
2.75 - 2.99	11.3	19.5	22.7	13.3
2.50 - 2.74	3.7	8.6	7.8	4.7
2.25 - 2.49	1.0	2.1	0.0	1.1
% Female	41.8	47.1	75.8	44.5
% Male	58.2	52.9	24.2	55.5
General Law Review				
Yes %	24.4	12.8	14.1	21.9
No %	71.9	82.3	80.5	74.1
Missing %	3.7	4.9	5.5	4.0
Specialty Law Review				
Yes %	26.0	22.4	26.6	25.4
No %	69.6	73.7	71.1	70.3
Missing %	4.5	3.9	2.3	4.3
Observations	1,826	384	128	2,338

This table uses data from the After the JD (AJD) representative survey of new lawyers in 2002-2003. Law school rankings are from the U.S. News and World Report in 2003. Schools in tiers 3 and 4 are not ordered within tier. Standard deviations in parentheses. Dollar figures are in 2002 dollars. See Appendix B1.1 for further details on this sample.

Table 2.2: Summary statistics by sector, University of Michigan Law School

	Private	Gov	Nonprofit	Overall
GPA 1st Year	0.1	0.1	0.1	0.1
	(0.9)	(1.2)	(0.9)	(0.9)
GPA Final	0.1	0.1	0.3	0.1
	(0.9)	(1.1)	(0.9)	(0.9)
Annual Wage	$137,929	$75,259	$56,791	$123,568
	(80,023)	(30,311)	(29,038)	(78,156)
Hourly Wage	$53.5	$31.8	$24.2	$48.5
	(27.2)	(11.2)	(12.4)	(26.9)
Weekly Hours	53.4	48.8	49.4	52.5
	(10.0)	(9.5)	(9.1)	(10.0)
Weeks Per Year	48.8	49.0	48.8	48.8
	(4.0)	(3.4)	(3.5)	(3.9)
Debt	$66,470	$60,266	$58,253	$65,046
	(44,725)	(46,138)	(42,330)	(44,735)
% Female	40.5	45.2	60.6	42.8
% Male	59.5	54.8	39.4	57.2
Law Journal %				
No	46.5	43.7	43.1	45.9
Yes	53.5	56.3	56.9	54.1
Observations	939	132	103	1,174

This table uses data from the University of Michigan Law School (UMLS) survey of its graduates, specifically, the 1996-2001 classes, 5 years after graduation, restricted to full-time workers (1000+ hours/year). GPAs are in standard deviation units around the population mean and come from administrative data. They do not average to 0 because higher GPA students respond more often. Dollar figures are in 2002 dollars. See Appendix B1.2 for further details on this sample.

Table 2.3: Sectoral choice probit marginal effects.

	(1)		(2)		(3)	
Rank 1-10	-0.030	(0.032)	-0.028	(0.029)	-0.010	(0.028)
Rank 11-20	-0.029	(0.028)	-0.019	(0.023)	-0.004	(0.022)
Rank 21-100	0.018	(0.018)	0.015	(0.016)	0.020	(0.015)
Tier 3	0.012	(0.020)	0.002	(0.019)	-0.002	(0.019)
GPA 4.00-3.75	0.077	(0.046)	0.049	(0.037)	0.037	(0.036)
GPA 3.74-3.50	0.079	(0.043)	0.033	(0.036)	0.028	(0.035)
GPA 3.49-3.25	0.065	(0.043)	0.034	(0.035)	0.023	(0.034)
GPA 3.24-3.00	0.082^{*}	(0.041)	0.052	(0.034)	0.044	(0.033)
GPA 2.99-2.75	0.013	(0.046)	0.004	(0.038)	0.003	(0.035)
Law Review General	0.025	(0.013)	0.016	(0.012)	0.011	(0.012)
Law Review Special	0.009	(0.015)	0.005	(0.014)	0.003	(0.014)
Male	0.072^{***}	(0.013)	0.038^{***}	(0.011)	0.031^{**}	(0.010)
Soc. Resp. Pref. 2			0.007	(0.013)	-0.014	(0.017)
Soc. Resp. Pref. 3			0.014	(0.012)	0.001	(0.011)
Soc. Resp. Pref. 4			0.002	(0.015)	-0.024	(0.022)
Soc. Resp. Pref. 5			-0.031	(0.023)	-0.088	(0.046)
Soc. Resp. Pref. 6			-0.109^{***}	(0.032)	-0.164^{**}	(0.058)
Soc. Resp. Pref. 7			-0.371^{***}	(0.047)	-0.342^{***}	(0.069)
Earn. Pot. Pref. 2					0.107^{*}	(0.050)
Earn. Pot. Pref. 3					0.143^{**}	(0.050)
Earn. Pot. Pref. 4					0.116^{*}	(0.050)
Earn. Pot. Pref. 5					0.146^{**}	(0.049)
Earn. Pot. Pref. 6					0.179^{***}	(0.050)
Earn. Pot. Pref. 7					0.177^{***}	(0.051)
Pseudo-R2	0.110		0.307		0.373	
N	1936		1936		1936	

This table uses data from the 2002-2003 AJD survey of new lawyers. Private Sector=1, Non-profit sector=0. Column 3 is our preferred specification. The preferences question in the AJD 2002-2003 survey are a set of questions of the form: "Thinking about the principal types of settings in which lawyers work (e.g., government, large law firms, business), how important was each of the following factors in determining the sector in which you began your professional career?" Answers range from 1="Not At All Important" to 7="Extremely Important". The prompts used in this table are worded, "Opportunity to do socially responsible work" and "Medium-to-long-term earning potential". Estimated effects are averaged over the sample and are relative to the omitted group for categorical variables. The omitted groups are: Tier 4 law school ranking, GPA 2.74-2.50, and responses of 1 for the preference questions. Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.4: Work preferences by sector (AJD 2002)

	Private	Gov	Nonprofit	All
Earnings potential	5.39	3.58	2.88	4.90
	(1.71)	(1.87)	(1.96)	(1.96)
Substantive interest	5.16	5.47	5.88	5.27
	(1.73)	(1.85)	(1.49)	(1.75)
Pay off debts	5.28	3.70	3.42	4.87
	(2.05)	(2.10)	(2.02)	(2.18)
Loan repayment assistance	1.89	2.66	3.65	2.16
	(1.62)	(2.05)	(2.49)	(1.85)
Develop specific skills	5.33	5.73	5.62	5.42
	(1.57)	(1.54)	(1.53)	(1.57)
Work-life balance	5.05	5.81	6.19	5.27
	(1.77)	(1.58)	(1.27)	(1.75)
Socially responsible	3.52	5.59	6.19	4.08
	(1.84)	(1.77)	(1.46)	(2.05)
Prestige	4.21	4.40	2.72	4.14
	(1.96)	(1.95)	(1.80)	(1.99)
Future career mobility	5.31	5.11	4.38	5.21
	(1.70)	(1.76)	(1.99)	(1.74)
N	1,106	264	99	1,469

The preferences question in the AJD 2002-2003 survey are a set of questions of the form: “Thinking about the principal types of settings in which lawyers work (e.g., government, large law firms, business), how important was each of the following factors in determining the sector in which you began your professional career?” Answers range from 1=“Not At All Important” to 7=“Extremely Important”. Standard deviations are in parentheses.

Table 2.5: Work preferences by law school rank

	1-10	11-20	21-100	Tier 3	Tier 4	All
Earnings potential	4.86	5.22	5.29	5.03	5.27	5.19
	(1.91)	(1.93)	(1.82)	(1.93)	(1.78)	(1.86)
Socially responsible	3.60	3.66	3.75	3.64	4.15	3.75
	(1.83)	(1.98)	(1.94)	(2.03)	(2.00)	(1.96)
N	125	174	551	204	146	1,200

The preferences question in the AJD 2002 survey are a set of questions of the form: "Thinking about the principal types of settings in which lawyers work (e.g., government, large law firms, business), how important was each of the following factors in determining the sector in which you began your professional career?" Answers range from 1="Not At All Important" to 7="Extremely Important". These two prompts are worded, "Medium-to-long-term earning potential" and "Opportunity to do socially responsible work". This table uses a pooled sample of private and non-profit lawyers. Standard deviations are in parentheses.

Table 2.6: AJD Private Sector Log Hourly Wage Regressions

	X_i		$X_i, Z_i =$ Social		$X_i, Z_i =$ Social, Earnings	
Rank 1-10	0.545***	(0.0498)	0.524***	(0.0494)	0.532***	(0.0489)
Rank 11-20	0.411***	(0.0484)	0.395***	(0.0478)	0.390***	(0.0479)
Rank 21-100	0.213***	(0.0403)	0.200***	(0.0401)	0.197***	(0.0400)
Tier 3	0.0795*	(0.0400)	0.0650	(0.0402)	0.0701	(0.0401)
GPA 4.00-3.75	0.501***	(0.137)	0.463***	(0.134)	0.458***	(0.128)
GPA 3.74-3.50	0.495***	(0.134)	0.453***	(0.131)	0.444***	(0.124)
GPA 3.49-3.25	0.400**	(0.133)	0.369**	(0.130)	0.361**	(0.123)
GPA 3.24-3.00	0.284*	(0.132)	0.256*	(0.129)	0.249*	(0.123)
GPA 2.99-2.75	0.177	(0.133)	0.165	(0.130)	0.178	(0.124)
GPA 2.74-2.50	0.201	(0.139)	0.178	(0.136)	0.193	(0.130)
Law Rev Gen	0.157***	(0.0257)	0.157***	(0.0253)	0.148***	(0.0250)
Law Rev Spec	0.0799**	(0.0262)	0.0931***	(0.0258)	0.0927***	(0.0256)
Male	0.0486*	(0.0208)	0.0346	(0.0208)	0.0285	(0.0206)
Soc. Resp. Pref. 2			0.0861	(0.0502)	0.0655	(0.0500)
Soc. Resp. Pref. 3			0.0490	(0.0453)	0.0369	(0.0455)
Soc. Resp. Pref. 4			0.0289	(0.0452)	0.00210	(0.0458)
Soc. Resp. Pref. 5			−0.0665	(0.0475)	−0.0838	(0.0474)
Soc. Resp. Pref. 6			−0.00939	(0.0540)	−0.0213	(0.0545)
Soc. Resp. Pref. 7			−0.209*	(0.0817)	−0.225**	(0.0781)
Earn. Pot. Pref. 2					0.119	(0.0848)
Earn. Pot. Pref. 3					0.138	(0.0870)
Earn. Pot. Pref. 4					0.132	(0.0788)
Earn. Pot. Pref. 5					0.182*	(0.0738)
Earn. Pot. Pref. 6					0.188*	(0.0733)
Earn. Pot. Pref. 7					0.287***	(0.0733)
Constant	2.967***	(0.137)	3.007***	(0.139)	2.841***	(0.150)
r2	0.335		0.356		0.372	
N	1826		1826		1826	

The preferences question in the AJD 2002-2003 survey are a set of questions of the form: "Thinking about the principal types of settings in which lawyers work (e.g., government, large law firms, business), how important was each of the following factors in determining the sector in which you began your professional career?" Answers range from 1="Not At All Important" to 7="Extremely Important". The prompts used in this table are worded, "Opportunity to do socially responsible work" and "Medium-to-long-term earning potential". The omitted groups are: Tier 4 law school ranking, GPA 2.49-2.25, and responses of 1 for the preference questions. Standard errors are in parentheses.
$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.7: AJD Nonprofit Sector Log Hourly Wage Regressions

	X_i	
Rank 1-10	-0.00802	(0.150)
Rank 11-20	0.0598	(0.0868)
Rank 21-100	0.0284	(0.0778)
Tier 3	0.172^{*}	(0.0787)
GPA 4.00-3.75	0.169	(0.158)
GPA 3.74-3.50	0.229	(0.143)
GPA 3.49-3.25	0.0770	(0.0871)
GPA 3.24-3.00	0.0595	(0.0854)
GPA 2.99-2.75	-0.00354	(0.0847)
Law Rev Gen	0.131	(0.120)
Law Rev Spec	0.0272	(0.0724)
Male	0.104	(0.0567)
Constant	2.589^{***}	(0.157)
r2	0.259	
N	128	

Standard errors are in parentheses.
$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.8: UMLS log hourly wage regressions by sector

	Private 5 years	Private 15 years	Nonprofit 5 years	Nonprofit 15 years
First year GPA	0.113^{***}	0.165^{***}	0.0486	0.0882
	(0.0151)	(0.0224)	(0.0368)	(0.0521)
Law Journal	0.0979^{***}	0.0780	0.121	0.00583
	(0.0289)	(0.0475)	(0.0696)	(0.120)
Male	0.0126	0.110^{*}	0.0171	-0.00753
	(0.0285)	(0.0455)	(0.0720)	(0.0992)
Class years after 1991	0.0632^{***}		-0.00376	
	(0.00621)		(0.0145)	
Class years after 1981		0.0138^{*}		0.0214
		(0.00659)		(0.0159)
Constant	3.424^{***}	4.205^{***}	3.011^{***}	3.501^{***}
	(0.0320)	(0.0503)	(0.0720)	(0.109)
r2	0.148	0.0804	0.0673	0.0491
N	1168	1015	109	116

The 5 year estimates use class years 1991-2000 and the 15 year estimates use class years 1981-1990. We restrict to full-time workers (1000+ hours/year). GPAs are in standard deviation units around the population mean and come from administrative data. Dollar figures are in 2002 dollars. Standard errors are in parentheses.
$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.9: Estimated average opportunity cost of working in the nonprofit sector

Model	
AJD	0.453^{***}
	(0.038)
UMLS 5 years after law school	0.640^{***}
	(0.039)
UMLS 15 years after law school	0.779^{***}
	(0.054)

The 5 and 15 year UMLS estimates use class years 1991-2000 and 1981-1990, respectively. Standard errors in parentheses. Standard errors are bootstrapped, using 200 repetitions.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.10: Estimated average opportunity cost of working in the nonprofit sector by law school rank

Law school rank	
All	0.453^{***}
	(0.033)
1-10	1.060^{***}
	(0.084)
11-20	0.737^{***}
	(0.065)
21-100	0.397^{***}
	(0.048)
Tier 3	0.161^{*}
	(0.067)
Tier 4	0.202^{***}
	(0.060)

Standard errors in parentheses. Standard errors are bootstrapped, using 200 repetitions. Estimates are in log hourly wages.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.11: Sensitivity of estimated average opportunity cost of working in the nonprofit sector to unobservables

ρ	AJD	UMLS
-0.5	0.841^{***}	1.102^{***}
	(0.034)	(0.049)
-0.4	0.760^{***}	1.006^{***}
	(0.034)	(0.044)
-0.3	0.682^{***}	0.913^{***}
	(0.033)	(0.041)
-0.2	0.605^{***}	0.821^{***}
	(0.033)	(0.038)
-0.1	0.528^{***}	0.730^{***}
	(0.033)	(0.038)
0	0.453^{***}	0.640^{***}
	(0.033)	(0.039)
0.1	0.377^{***}	0.550^{***}
	(0.033)	(0.042)
0.2	0.301^{***}	0.459^{***}
	(0.034)	(0.046)
0.3	0.225^{***}	0.365^{***}
	(0.035)	(0.052)
0.4	0.147^{***}	0.269^{***}
	(0.035)	(0.060)
0.5	0.067	0.168^{*}
	(0.037)	(0.069)

The UMLS column is based on the 5-year estimates as shown in Table 2.9. Standard errors in parentheses. Standard errors are bootstrapped, using 200 repetitions. Estimates are in log hourly wages. ρ represent the correlation between the sector selection and wage equations.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.12: Sector Switchers from AJD

2002/2007 sector	Nonprof/Nonprof	Nonprof/Priv	Priv/Priv	Priv/Nonprof
Log hourly wage 2002	3.02	3.12	3.59	3.48
	(0.28)	(0.31)	(0.45)	(0.68)
Log hourly wage 2007	3.25	3.92	3.98	3.53
	(0.30)	(0.77)	(0.55)	(0.40)
Social Resp. Preference	6.31	5.71	3.41	3.69
	(1.31)	(1.65)	(1.78)	(2.24)
Earning Potential Preference	2.38	3.76	5.45	5.07
	(1.70)	(2.05)	(1.70)	(1.87)
Observations	32	17	551	29

This table uses data from the 2002-2003 and 2007-2008 AJD surveys of new lawyers. Wages are in 2002 dollars. Standard deviations in parentheses.

Table 2.13: AJD switchers: log hourly wage regression

	Sector in Wave 1	
	Nonprofit	Private
	(1)	(2)
Switched Sector	0.567***	−0.340***
	(0.143)	(0.088)
Sector Switchers	17	29
Observations	98	1,160
R^2	0.792	0.819

Notes: This table uses panel data from Waves 1 and 2 of AJD. It includes individual and wave fixed effects. Each observation is a log hourly wage of an individual in either Wave 1 (about 2 years after law school) or Wave 2 (about 7 years after law school). Standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.14: Sector Switchers from UMLS

A. Wages and GPA by Sector at Years 1 and 5 After Law School (UMLS Class Years 1991-2000)				
Year1/Year4 sector	Nonprof/Nonprof	Nonprof/Priv	Priv/Priv	Priv/Nonprof
Log annual wage (year 1)	10.43	10.34	11.17	10.94
	(0.28)	(0.39)	(0.40)	(0.64)
Log annual wage (year 4)	10.64	10.83	11.58	10.81
	(0.50)	(0.57)	(0.49)	(0.53)
GPA 1st Year	0.04	-0.70	0.20	0.16
	(0.85)	(1.03)	(0.90)	(0.93)
GPA Final	0.13	-0.72	0.20	0.13
	(0.83)	(0.97)	(0.88)	(0.81)
Observations	42	19	1124	56

B. Wages and GPA by Sector at Years 1 and 15 After Law School (UMLS Class Years 1981-1990)				
Year1/Year14 sector	Nonprof/Nonprof	Nonprof/Priv	Priv/Priv	Priv/Nonprof
Log annual wage (year 1)	10.46	10.25	11.13	11.00
	(0.27)	(0.93)	(0.33)	(0.41)
Log annual wage (year 14)	10.95	11.42	12.15	11.27
	(0.49)	(0.69)	(0.75)	(0.83)
GPA 1st Year	-0.36	-0.67	0.18	0.42
	(0.86)	(0.91)	(0.89)	(1.11)
GPA Final	-0.37	-0.59	0.19	0.46
	(1.03)	(0.94)	(0.88)	(1.13)
Observations	18	14	853	73

This Table shows wages and GPAs by sector switchers grouping from the University of Michigan Law School. Panel A shows switchers from their first year to their fourth year after law school. Panel B shows switchers from their first year to their fourteenth year after law school. Wages are in 2002 dollars. GPAs are in standard deviation units around the population mean and come from administrative data. They do not average to 0 because higher GPA students respond more often. Standard deviations are in parentheses.

Table 2.15: UMLS switchers: log annual wage regressions

	Sector in year 1			
	Nonprofit	Private	Nonprofit	Private
	(1)	(2)	(3)	(4)
Switched Sector	0.278	-0.537^{***}	0.684^{*}	-0.748^{***}
	(0.152)	(0.071)	(0.301)	(0.091)
Sector Switchers	19	56	14	73
Initial Earnings Year	1	1	1	1
Later Earnings Year	5	5	15	15
Class Years	1991-2000	1991-2000	1981-1990	1981-1990
Observations	122	2,360	64	1,852
R^2	0.648	0.734	0.693	0.769

Notes: This table uses panel data from UMLS. It includes individual and survey year fixed effects. Each observation is a log annual wage of an individual in either the first and fifth years after law school (columns 1 and 2) or the first and fifteenth year after law school (columns 3 and 4). We use annual rather than hourly wages since hours of work are not available in the first year after law school. Columns 1 and 2 use class years 1991-2000 and columns 3 and 4 use class years 1981-1990. Standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.16: Sector Plan Switchers from the UMLS

Entering/Leaving Plans	Nonprof/Nonprof	Nonprof/Priv	Priv/Priv	Priv/Nonprof	All
Annual Wage (year 1)	\$51,017	\$74,715	\$71,777	\$48,053	\$68,849
	(31,306)	(22,066)	(20,901)	(20,449)	(23,869)
Annual Wage (year 4)	\$70,056	\$83,452	\$96,345	\$74,696	\$91,126
	(56,933)	(33,727)	(35,552)	(33,445)	(39,791)
Annual Wage (year 14)	\$159,204	\$133,392	\$244,586	\$152,577	\$219,978
	(258,548)	(120,726)	(258,428)	(130,053)	(248,751)
Year 1 GPA	0.30	-0.04	0.16	0.21	0.16
	(1.10)	(0.82)	(0.90)	(1.23)	(0.93)
Final GPA	0.42	0.02	0.17	0.21	0.19
	(1.04)	(0.80)	(0.89)	(1.23)	(0.92)
Observations	86	69	499	16	670

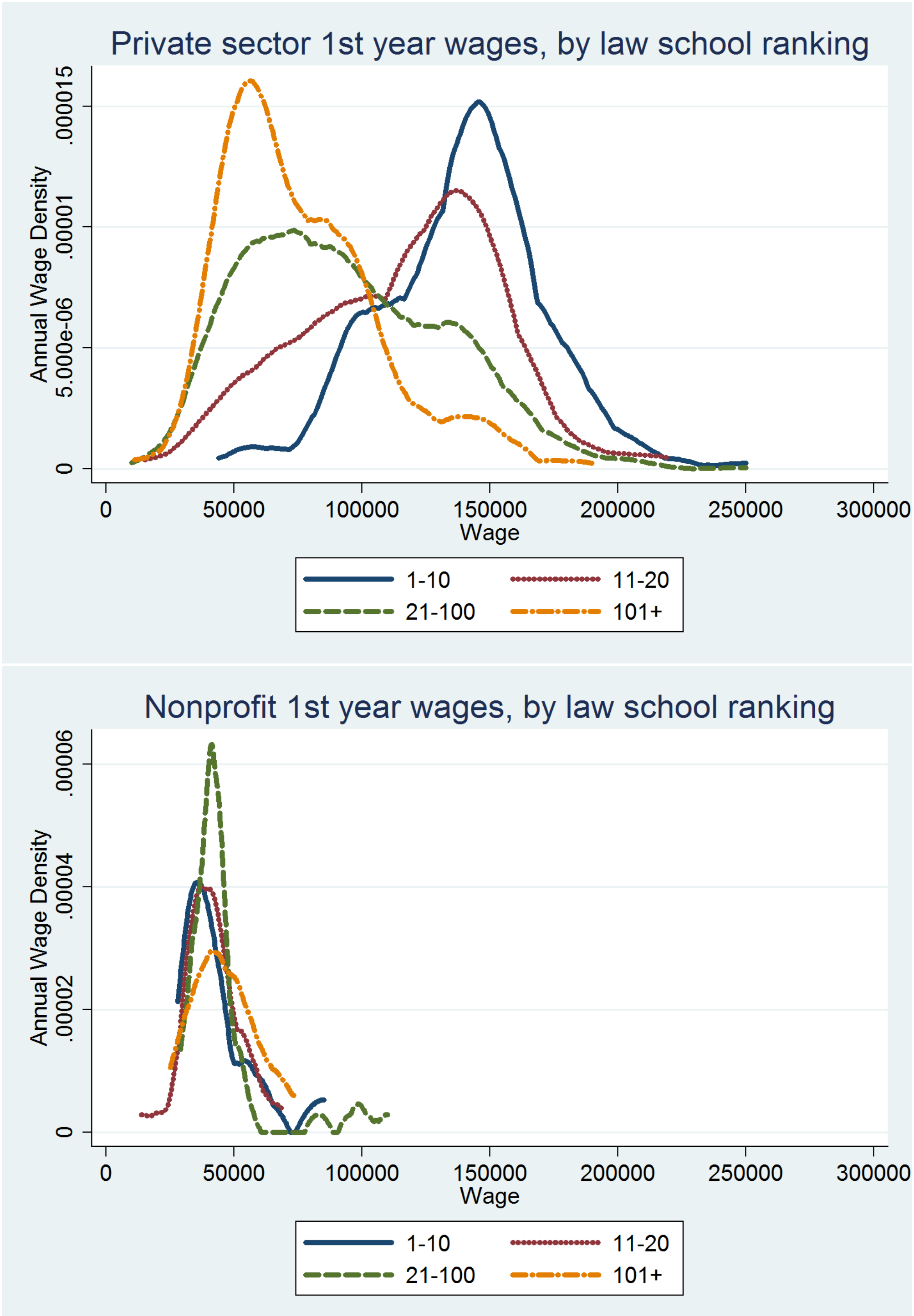
This Table shows wages and GPAs by planned long-term career sector plan switchers grouping, when entering and leaving law school, for University of Michigan Law School, class years 1992-1991. Wages are in 2002 dollars. GPAs are in standard deviation units around the population mean and come from administrative data. They do not average to 0 because higher GPA students respond more often. Standard deviations are in parentheses.

Table 2.17: Debt by Law School Rank

	1-10	11-20	21-100	Tier 3	Tier 4	All
Debt	$73,848	$63,291	$60,614	$68,078	$71,062	$64,874
	(48,921)	(38,562)	(39,354)	(40,762)	(35,272)	(40,350)
Observations	194	250	904	326	236	1,910

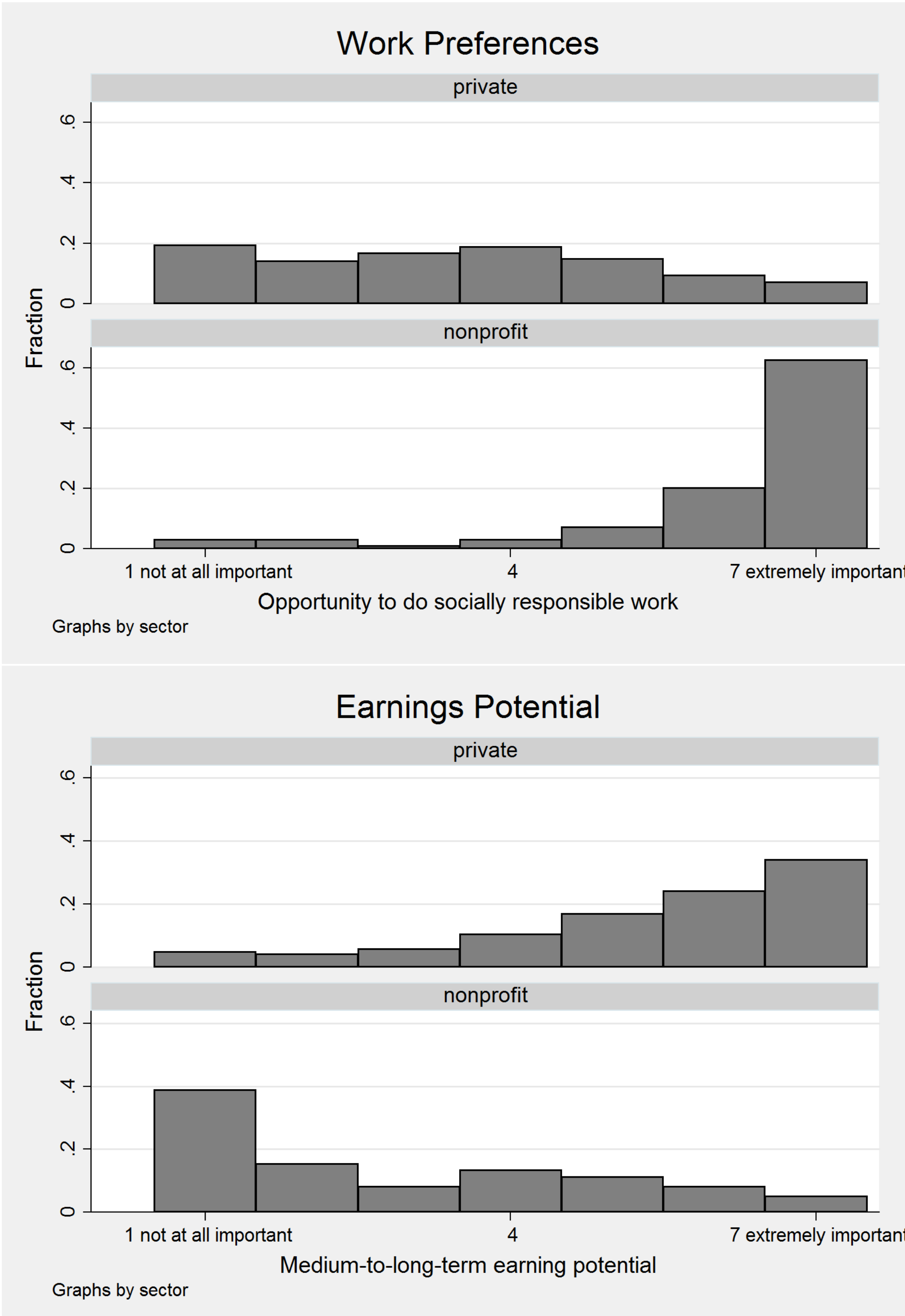
This table uses data from the 2002-2003 AJD survey of new lawyers. Standard deviations in parentheses. Dollar figures are in 2002 dollars.

Figure 2.1: Distribution of Private Sector Annual Compensation by Law School Rank in the Private and Nonprofit Sectors



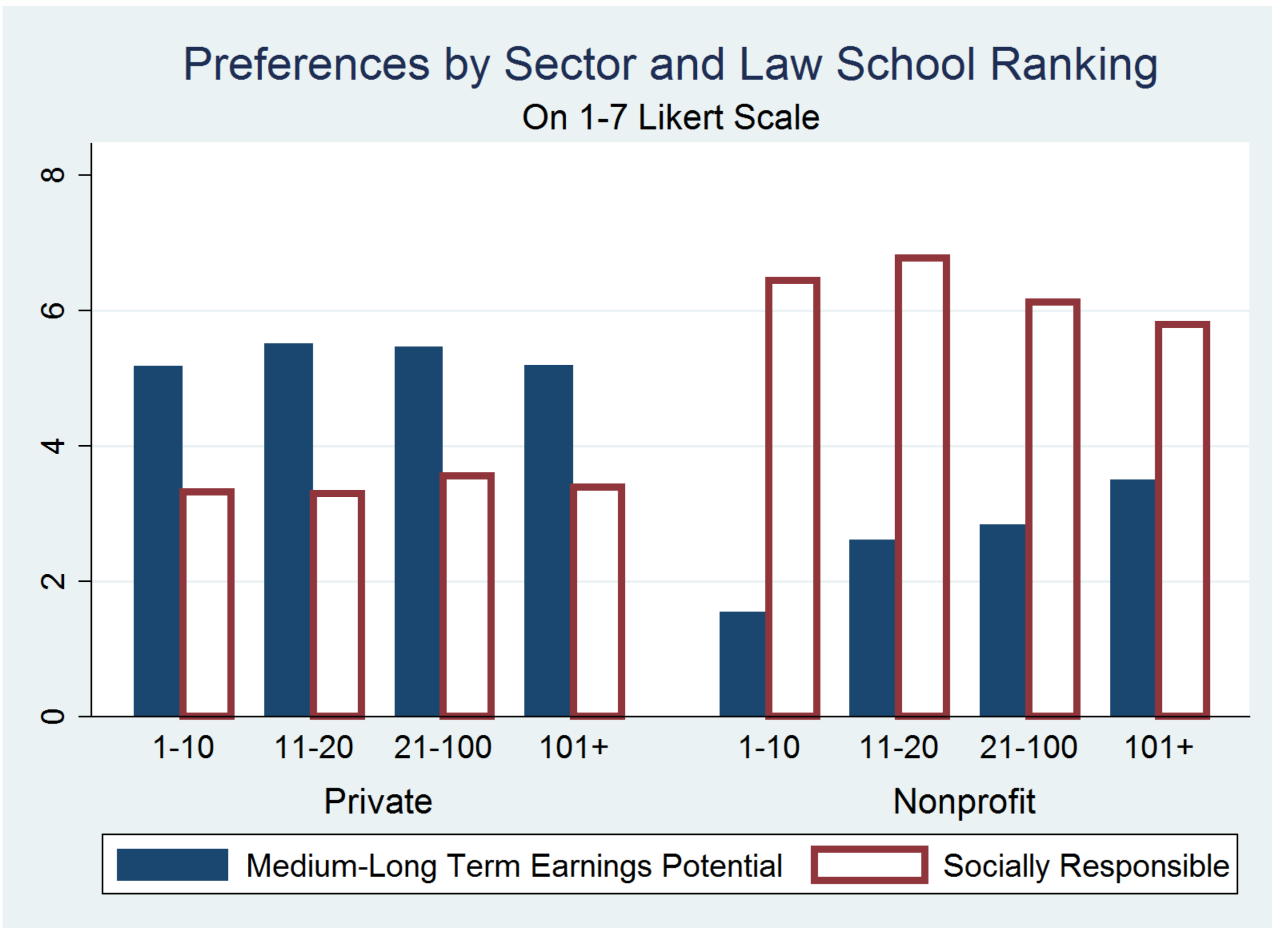
This figure uses data from the AJD survey of new lawyers in 2002-2003.

Figure 2.2: Preferences and Sector Choice



This figure uses data from the AJD survey of new lawyers in 2002-2003.

Figure 2.3: Preferences and Skill



This figure uses data from the AJD survey of new lawyers in 2002-2003.

Figure 2.4: Conceptual Diagram of Sorting by Skill and Preferences

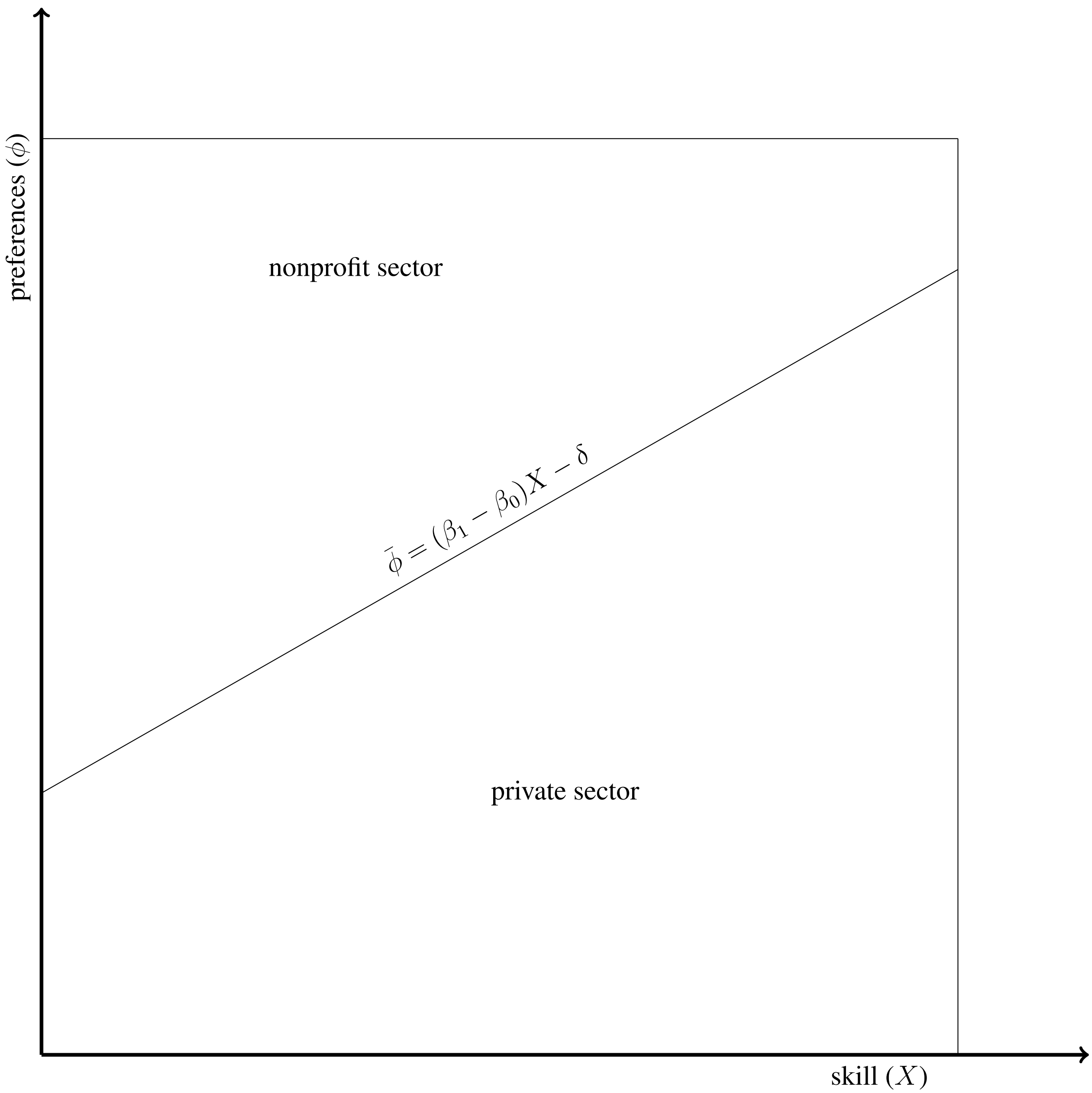
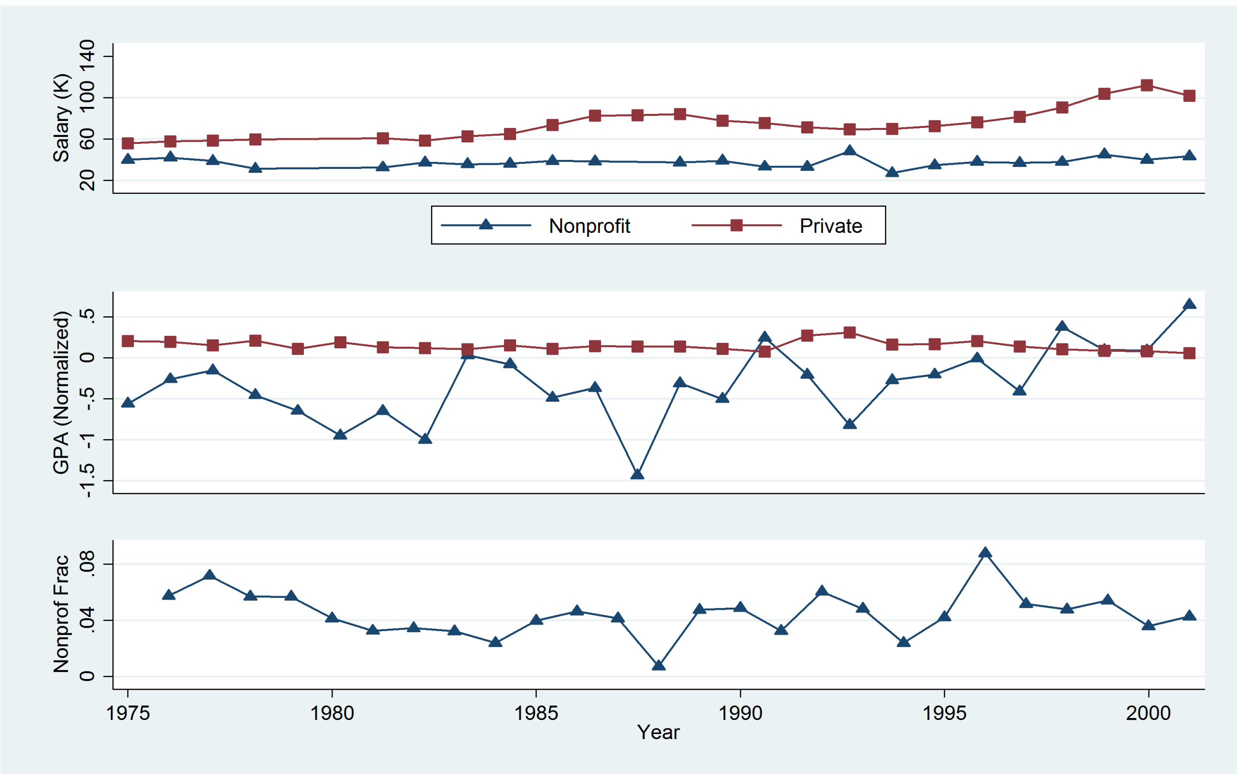


Figure 2.5: Lawyer Pay, GPA and Sectoral Sorting Over Time



This figure uses data from the University of Michigan Law School Survey. Annual salary is in 2002 dollars. GPA's are within year standard deviations away from the mean.

B1 Detailed Data Description

B1.1 After the JD

Survey Design and Response Rates

We use waves 1 and 2 of the restricted version of the After the JD (AJD) panel sample of lawyers who first passed the bar in 2000 and graduated from law school between 1998 and 2000. 95% of respondents graduated in 2000. AJD waves 1 and 2 were conducted in 2002-2003 and 2007-2008 respectively by the American Bar Foundation (ABF), with a response rate of 71% in wave 1 and 51% in wave 2. Wave 1 had 4,538 respondents who met the above sampling criteria, while wave 2 had 3,704, with a total of 5,353 lawyers who responded to at least one wave.

This response rate is for sample members who the survey administrators located and who met the criteria for the study. In wave 1, the total response rate of all lawyers in the entire sample, including those later not located or who did not meet the sample's criteria was 50%. About 20% could not be located. By wave 2, members who did not meet the sampling criteria were excluded and ABF located and surveyed 98% of the entire sample of eligible respondents, including those that did not reply during wave 1. 27% of repondents in wave 2 did not take the survey in wave 1. Not all respondents answered each question.

The survey is nationally representative with a stratified sample design based on surveying 18 legal markets, including the largest four (New York City, the District of Columbia, Chicago and Los Angeles), as well as 14 other areas consisting of smaller metropolitan areas or entire states. We use the AJD sampling weights in all regressions.

Sample Restrictions

In all of our empirical results, we restrict our sample of the 5,353 individuals who responded to at least one of the two waves to 2,338 individuals as follows. We drop individuals in the following groups: lowest GPA category, did not take graded courses, did not report a GPA, law degree is from outside the US, law degree is from an unaccredited school, annual wage is greater than $250,000, did not report an annual wage, did not report weekly work hours, reported "other" as their work sector, did not report their work sector and age over 40. Further restricting our sample to only private or nonprofit lawyers (removing government lawyers) reduces our sample to 1,954. Some tables in our analysis have smaller sample sizes because they rely on additional survey questions which not all respondents answered.

We restrict the sample to include respondents who completed key questions (including GPA). However, there is still some nonresponse to some of our other controls. Because all of our control variables are categorical, for each question we create an additional category of nonrespondents. Thus, the regressions capture the conditional expectation in the population that answered the question.

Variable Construction

The key variables are the sector in which the lawyer is employed and annual income. We also describe all of the controls for skill that are used in the regressions. The preference variables are described in detail in the paper.

Sector: We use a question about the type of organization in which the lawyer works to group lawyers into sectors. The question asks, "What type of organization is [your current primary employer]?" AJD provides respondents with 13 choices. Table B1 shows how we group the 13 possible responses into our 3 sectors, private, government and nonprofit.

Income and Hours: Our measure of income is pre-tax income including bonuses. We use responses to the following question: "What is your total annual salary (before taxes) including estimated bonus, if applicable, at your current job?" For our measure of hours we use a measure of "typical" hours: "How many hours are you expected to work during a typical week at your job?" This question is distinct from the number of $billable$ hours in a week.

Constructing Hourly Wage: Data on number of weeks worked per year is not available. We construct an hourly wage variable assuming that lawyers in all sectors worked 50 weeks a year. We chose 50 because it is both the mode and mean of responses to the UMLS survey question about weeks worked per year, for lawyers in class years 1996-2001 who worked at least 40 hours per week. One concern with this procedure is if hours and hourly wage are positively correlated, then we might not be accounting for the bundling of hours and salary and as a result we would be overstating the compensation of private sector lawyers because they consume less leisure. However, hours and hourly wage are weakly negatively correlated; with a correlation coefficient of $-.20$. We report all wages in real (2002) dollars.

Educational Background: All regressions in the paper include the following variables about law school:

- Categorical variables for law school GPA in 0.25 GPA unit bins;
- Law school tier;
- Whether on a specialty law review;
- Whether on a general law review;

We also include the following variables about a lawyer's undergraduate education:

- Whether attended a public or a private school;
- Undergraduate GPA in 0.25 unit bins;
- Whether graduated in the top 10% of the undergraduate class;
- Whether majored in science as an undergraduate.

Demographics We include an indicator variable for gender, and a categorical variable for race/ethnicity.

A Note on Variables in Regressions Using AJD Data

Our wage and sector regressions, for conciseness, all include unreported coefficients on the above undergraduate variables, and on indicators for missing values for all categorical variables, where applicable.

B1.2 University of Michigan Law School Survey

Survey Design and Response Rates

The University of Michigan Law School (UMLS) panel dataset contains survey responses from UMLS students who graduated between the years of 1952-2001. The law school conducted surveys each year and surveyed lawyers as often as 5, 15, 25, 35 and 45 years after graduation, although not all class years were surveyed at each of these milestones, a limitation which restricts our analysis in some cases. Table B2 shows the years in which the survey was conducted after graduation, by class year. The most recent survey year available to us is 2006. Most of the questions did not change substantially over the years and respondents were asked both about the present day and their time in law school. The dataset contains a total of 16,921 individuals (including non-respondents), and 34,828 unique opportunities to respond, with a response rate of about 66% each year. We typically pool responses across 5 or more years in our analysis, because there are only around 200 respondents per class year and only a small fraction of them are in the nonprofit sector.

Variable Construction

Sectoral Plans:
We use the following variables:

- v0110: employment plans when entering law school;
- v0111: employment plans when leaving law school;

Sector:
We use the following questions to code sector:

- v0444, v1429, v2429: sector 1, 5 and 15 years after graduation, respectively.

Each question has a slightly different set of responses. UMLS does not use the same sectoral categories as AJD, or across relevant questions. For post-law school employment questions, we consider responses of "Legal service" and "Public interest" to be the nonprofit sector, while for questions about plans during law school we count the “Legal Services” and “Teacher” categories as nonprofit, noting that “Public interest” is not one of the categories in this case.

For v0444:

- Private: Private firm or business;
- Government: Government;
- Nonprofit: legal service or public interest.

For v1429 and v2429:

- Private: Private firm; Fortune 500; Other business; banking/finance; accounting firm; insurance.
- Government: Federal government; state or local government;
- Nonprofit: legal service; public interest; education.

Income and Hours:

We use the following variables for income:

- v1851: income in first year after law school;
- v1852: income in fourth year after law school
- v2853: income in fourteenth year after law school
- v3854: income in twenty four years after law school
- v4855: income in thirty four years after law school

We report all wages in real (2002) dollars. To construct the present discounted value of lifetime earnings we smooth between these reports and discount annual earnings at a 0.95 rate.

We use the following variables for hours and weeks:

- v1479 and v2479: hours worked per week five and fifteen years after law school;

- v1480 and v2480: weeks worked per year five and fifteen years after law school.

Educational Background:

We include the following variables:

- v0871: first year GPA in standard deviation units;
- v0872: final GPA in standard deviation units;
- v0143: law journal (any)

Table B1: Construction of the Sector Variable (AJD 2002)

Detailed Sector	Number	Sector	Number
Solo Practice	61	Private	1,826
Private Law firm	1,653		
Prof Serv Firm	24		
Other Fortune 1000	35		
Other business/indus	53		
Fed (including clerk)	115	Gov	384
State/Local (including clerk)	269		
Legal Service/Public Defender	78	Nonprofit	128
public interest	28		
other nonprofit	13		
educational instit	7		
labor trade union	2		

Table B2: UMLS Survey Years

	Years After Graduation				
Class Year	5	15	25	35	45
1952-1961	no	yes	no	no	yes
1962-1967	no	yes	no	yes	no
1968-1971	yes	yes	no	yes	no
1972-1981	yes	yes	yes	no	no
1982-1991	yes	yes	no	no	no
1992-2001	yes	no	no	no	no

Table B3: Law School Financial Support Percentages by Sector

	Private	Gov	Nonprofit
employment	14.1	14.1	9.8
federal loans	40.3	43.7	48.1
private loans	6.3	7.1	7.3
other loans	5.2	4.8	4.6
school based grants	8.1	7.8	7.8
other grants	0.6	0.4	0.0
spouse/partner	4.5	3.3	6.8
parent or relative	17.7	14.6	11.4
previous savings	2.7	3.8	4.0
veterans benefits	0.5	0.8	0.2
other	3.7	1.2	5.6
Observations	291	83	36

Table B4: Parents' Education by Sector, AJD

	sector			
	private	gov	nonprofit	Total
	%	%	%	%
father's education				
grade school	3.4	1.2	6.5	3.2
some high school	2.5	2.4	3.3	2.5
high school diploma or equivalent	15.4	16.3	10.6	15.2
trade or vocational school	4.4	3.6	5.7	4.4
associate or two-year degree	10.8	12.5	16.3	11.5
bachelors or four-year degree	20.3	16.6	20.3	19.6
law degree (j.d.)	8.7	10.1	8.1	9.0
some graduate or professional school	4.4	3.3	1.6	4.0
graduate or professional degree	30.1	34.1	27.6	30.7
mother's education				
grade school	2.9	1.5	4.8	2.8
some high school	3.0	1.8	2.4	2.7
high school diploma or equivalent	21.6	18.5	22.4	21.1
trade or vocational school	4.6	5.7	6.4	4.9
associate or two-year degree	16.7	19.9	15.2	17.2
bachelors or four-year degree	25.0	25.6	20.0	24.8
law degree (j.d.)	1.5	1.5	0.8	1.5
some graduate or professional school	4.8	3.9	3.2	4.5
graduate or professional degree	19.9	21.7	24.8	20.6
N	1,320	336	125	1,781

CHAPTER 3

Recruitment of Foreigners in the Market for Computer Scientists in the US

3.1 Introduction

An increasingly high proportion of the scientists and engineers in the US were born abroad. At a very general level, the issues that come up in the discussion of high skilled immigration mirror the discussion of low skilled immigration. The most basic economic arguments suggest that both high-skill and low-skill immigrants: (1) impart benefits to employers, to owners of other inputs used in production such as capital, and to consumers, and (2) potentially, impose some costs on workers who are close substitutes (Borjas (1999)). On the other hand, the magnitude of these costs may be substantially mitigated if US high skilled workers have good alternatives to working in sectors most impacted by immigrants (Peri and Sparber (2011), Peri et al. (2013)). Additionally, unlike low skilled immigrants, high skilled immigrants contribute to the generation of knowledge and productivity through patenting and innovation. Doing so both serves to shift out the production possibility frontier in the US and may also slow the erosion of the US comparative advantage in high tech (Freeman (2006); Krugman (1979)).

In this paper we study the impact of high skilled immigration on the labor market for computer scientists (CS) in the US, during the Internet boom of the 1990s, and the subsequent slump in the early 2000s. During this period, we observe a substantial increase in the number of temporary non-immigrant visas awarded to high skilled workers, and individuals with computer-related occupations becoming the largest share of H-1B visa holders (US General Accounting Office, 2000). Given these circumstances, it is of considerable interest to investigate how the influx of foreigners affected the labor market outcomes for US computer scientists during this period.

In order to evaluate the impact of immigration on CS domestic workers, we construct a dynamic model that characterizes the labor supply and demand for CS workers during this period. We build into the model the key assumption that labor demand shocks, such as the one created by the dissemination of the Internet, can be accommodated by three sources of CS workers: recent college graduates with CS degrees, US residents in different occupations who switch to CS jobs, and skilled foreigners. Furthermore, firms face a trade-off when deciding to employ immigrants: foreigners are potentially either more productive or less costly than US workers, but there are extra recruitment costs associated with hiring them.

The approach we take in this paper is distinctly partial equilibrium in nature – we focus on the market for computer scientists and ignore any wider impacts that high skilled immigration might have on the U.S. economy (Nathan (2013)). While we believe this approach can potentially be used to understand the impact that the availability of high skilled foreign labor might have had for this market, this approach precludes any analysis of the overall welfare impact of the H-1B program in particular or high skilled immigration more generally.

The predictions of the model on the impacts of immigration on wages depend on the elasticity of labor demand for computer scientists. As long as the demand curve slopes downwards, the increased availability of foreign computer scientists will put downward pressure on the wages for computer scientists in the US. However, as we discuss further in Section 4.4, there are a number of considerations that might lead us to think otherwise in the case of computer scientists. First, even in a closed economy, the fact that computer scientists contribute to innovation reduces the negative effects foreign computer scientists might have on the labor market opportunities for skilled domestic workers. In addition, in an increasingly global world, we might expect that restrictions on the hiring of foreign skilled workers in the US would lead employers to increase the extent to which they outsource work. Indeed, if computer scientists are a sufficient spur to innovation, or if it is easy for domestic employers to offshore work, any negative effects that an increase in the number of foreign computer scientists working in the US might have on the domestic skilled workforce would be completely offset by increases in the domestic demand for computer scientists. In the end this issue comes down to the slope of the demand curve for computer scientists.[1]

We use data on wages, domestic and foreign employment, and undergraduate degree completions by major, during the late 1990s and early 2000s to calibrate the parameters of our model such that it reproduces the stylized facts of the CS market during the period. Next, we use the calibrated model to simulate counterfactuals on how the economy would behave if firms had a restriction on the number of foreigners they could hire. Conditional on our assumptions about the slope of the demand curve for computer scientists, our simulation suggests that had US firms not been able to increase their employment of foreign computer scientists above its 1994 level, CS wages would be 2.8-3.8% higher in 2004. Furthermore, the number of Americans working in the CS industry would be 7.0-13.6% higher, the total number of CS workers would be 3.8-9.0% lower and the enrollment levels in computer science would be 19.9-25.5% higher than the observed levels in 2004.

Within the confines of the model, the predictions of our model do not depend on the specific choice we made for non-calibrated parameters, with one important exception. The exception: crowd out in the market for computer scientists depends crucially on the elasticity of demand for their services. Ideally, we would be able to use exogenous supply shifts to identify the slope of the demand curve for computer scientists, while we use exogenous shifts in demand to identify supply curves. We believe that largely exogenous technological breakthroughs in the 1990s increased the demand for computer scientists, allowing us to identify supply curves.[2] In other contexts, researchers have treated the increase in foreign born workers in the US economy as exogenous. However, in the current context, immigration law in the US implies that most of the foreign born and trained individuals who migrate to the US to work as computer scientists do so because they

[1] In this discussion we are assuming that foreign trained computer scientists are close substitutes for domestically trained ones. If foreign and domestically trained computer scientists are imperfect substitutes for each other, then the impact that the increased immigration will have on domestically trained computer scientists will also depend on the degree of substitutability between computer scientists trained domestically and abroad.

[2] These include the introduction of the World Wide Web, web browsers, and of search engines. During this time, Microsoft developed popular user-friendly operating systems, and Linux and other free and open-source software packages grew to power much of the Internet's server infrastructure. Sun Microsystems introduced the Java programming language and various service providers made e-mail available to a wider base of consumers. These types of software innovation, along with steady, rapid improvements to computer hardware and reductions in its cost permanently changed the structure and nature of the industry.

are sponsored by US based firms. Thus, it seems implausible to treat the number of foreign born computer scientists in the US as an exogenous increase in supply. In the end, without credible sources of identifying information, we resort to parametrically varying the elasticity of the demand for computer scientists through, what we will argue is a plausible range, from -1.3 to -4.0.

This paper constitutes a contribution to two different dimensions of the research literature. First, our study can be seen as an extension of the models of the market for scientists and engineers developed by Freeman (1975, 1976) in the 1970s and refined by Ryoo and Rosen (2004a) more recently. In Ryoo and Rosen's model, employers are restricted to hiring recent graduates from US engineering programs. In our model, employers can also hire both foreigners and US based individuals not trained as computer scientists. As a result, the supply of CS workers implied by our model is substantially more elastic than implied by the Ryoo and Rosen model, especially in the short term. More importantly, the substantial number of skilled foreign workers affects how the labor and education markets adjusts to an increase in the demand for skilled labor. Second, our paper relates to the recent literature on the potential impact that the hiring of high skilled immigrants might have on the wages and employment prospects of US natives.

We review this literature in detail, and describe the market for CS workers in section 2. Section 3 presents the dynamic model we build to characterize the market for CS workers when firms can recruit foreigners. In section 4, we describe how we calibrate the parameters of the model and the counterfactual simulations where firms have restrictions on the number of foreigners they can hire. We conclude with section 5 which presents a discussion based on the results of the paper.

3.2 The Market for Computer Scientists in the 1990s

3.2.1 The Information Technology Boom of the Late 1990s

During the mid 1990s, we observe the beginning of the utilization of the Internet for commercial purposes in the United States[3] and a substantial increase in the number of Internet users. One

[3]The decommissioning of the National Science Foundation Network in April of 1995 is considered the milestone for introducing nationwide commercial traffic on the Internet. (Leiner et al. (1997)).

indicator of a contemporaneous change in demand for IT workers is the rise of R&D expenditure of firms in both the computer programming services, and the computer related equipment sector. Specifically, the share of total private R&D of the firms of these two industries increased from 19.5% to 22.1% between 1991 and 1998 (author's computations using Compustat data). The entry and then extraordinary appreciation of tech firms like Yahoo, Amazon and eBay provides a further testament to the "boom" in the IT sector prior to 2001.

These technological innovations had a dramatic effect on the labor market for computer scientists. According to the Census, the number of employed individuals working either as computer scientists or computer software developers (CS) increased by 161% between the years 1990 and 2000. As a comparison, during the same period, the total number of employed workers with at least a bachelor degree increased by 27%, while the number of workers in other STEM occupations increased by 14%.[4] Table 3.1 shows computer scientists as a share of the college educated workforce and the college educated STEM workforce. In each case, the share was rising before 1990, but rises dramatically during the 1990s. Indeed, by 2000 more than half of all STEM workers are computer scientists. In Figure 3.1a, we use the CPS to show a similar pattern, additionally showing that the growth of CS employment started in the second half of the decade - the same period as the dissemination of the Internet. There is no doubt this was a period of employment expansion of the CS workforce.

On top of employment decisions, there is evidence that Internet innovation also affected educational choices of students. We show in Figure 3.1b that the number of bachelor degrees awarded in computer science as a fraction of both the total number of bachelor degrees and the number of STEM major degrees increased dramatically during this period. The CS share of total bachelor degrees increased from about 2% in 1995 to more than 4% in 2002. Even when compared to other STEM majors, it is clear from the figure that for college students, the decision to study computer science also responded to the Internet boom.

In addition to affecting employment and enrollment decisions, there is also empirical evidence

[4]Here and elsewhere our tabulations restrict the analysis to workers with at least a bachelor degree and use the IPUMS suggested occupational crosswalk. Other STEM occupations are defined as engineers, mathematical and natural scientists.

that computer scientist wages responded to expanding Internet use. From the Census, we observe a 18% increase in the median real weekly wages of CS workers between 1990 and 2000. The CPS presents similar patterns: starting in the year 1994 we observe in Figure 3.1c that wages of computer scientists increased considerably when compared to both workers with other STEM occupations and all workers with a bachelor degree. In fact, while during the beginning of the 1990s, the earnings of CS workers were systematically lower than other STEM occupations, the wage differential tends to disappear after 1998.[5]

3.2.2 The Immigrant Contribution to the Growth of the High Tech Workforce

Employment adjustments in the market for computer scientists happened disproportionately among foreigners during the Internet boom. Evidence for this claim is found in Table 3.1 and Figure 3.1d, where we use the Census and CPS to compare the share of foreign computer scientists to the share of foreign workers in other occupations.[6] In the second half of 1990s, the foreign fraction of CS workers increased considerably more than both the foreign fraction of all workers with a bachelor degree and the foreign fraction of all workers in a STEM occupation. In particular, foreigners were less represented among individuals working as computer scientists than in other STEM occupations in 1994. However, with the dissemination of the Internet in the later years of the decade, foreigners became a more important part of the pool of CS workers, as foreigners comprised 29.6% of the increase in CS workers.

The growth in the representation of the foreign born among the US computer scientist workforce was fueled by two developments. First, there was a truly dramatic increase in the foreign supply of men and women with college educations in science and engineering fields (Freeman (2009)). To take one important example, in India, the number of first degrees conferred in science and engineering rose from 176 thousand in 1990 to 455 thousand in 2000. Second, the Immigration

[5]It seems likely to us that wages increased as well for complementary jobs to computer scientists, such as marketing and sales staff at software companies. But we leave such spillovers for later research.

[6]Here and elsewhere, we define foreigners as who immigrated to the US after the age of 18. We believe that this definition is a proxy for workers who arrived to the US with non-immigrant visa status.

Act of 1990 established the H-1B visa program for temporary workers in "specialty occupations."[7] The regulations define a "specialty occupation" as requiring theoretical and practical application of a body of highly specialized knowledge in a field of human endeavor including, but not limited to, architecture, engineering, mathematics, physical sciences, social sciences, medicine and health, education, law, accounting, business specialties, theology, and the arts. In addition, applicants are required to have attained a bachelor's degree or its equivalent as a minimum.

Firms that wish to hire foreigners on H-1B visas must first file a Labor Condition Application (LCA). In LCA's for H-1B workers, the employer must attest that the firm will pay the non-immigrant the greater of the actual compensation paid to other employees in the same job or the prevailing compensation for that occupation, and the firm will provide working conditions for the non-immigrant that do not cause the working conditions of the other employees to be adversely affected. At that point, prospective H-1B non-immigrants must demonstrate to the US Citizenship and Immigration Services Bureau (USCIS) in the Department of Homeland Security (DHS) that they have the requisite education and work experience for the posted positions. USCIS then may approve the petition for the H-1B non-immigrant for a period up to three years. The visa may be extended for an additional three years, thus a foreigner can stay a maximum of six years on an H-1B visa, though firms can sponsor H-1B visa holders for a permanent resident visa. An important feature of the H-1B visa is that the visa is for work at the specific firm. As a result, workers are effectively tied to their sponsoring firm.

Since 1990 there has been a cap in the number of H-1B visas that can be issued. Initially this cap was set at 65,000 visas per year. In the initial years of the program, the cap was never reached, By the mid-1990s, however, the allocation tended to fill each year on a first come, first served basis, resulting in frequent denials or delays on H-1Bs because the annual cap had been reached. After lobbying by the industry, at the end of the decade, Congress acted to raise the cap first to 115,000

[7] The Immigration and Nationality Act of 1952 established the precursor to the H-1B visa, the H-1. The H-1 non-immigrant visa was targeted at aliens of "distinguished merit and ability" who were filling positions that were temporary. Nonimmigrants on H-1 visas had to maintain a foreign residence. The Immigration Act of 1990 established the main features of H-1B visa as it is known today, replacing "distinguished merit and ability" with the "specialty occupation" definition. It also dropped the foreign residence requirement and also added a dual intent provision, allowing workers to potentially transfer from an H-1B visa to immigrant status.

for FY1999 and to 195,000 for FY2000-2003. The cap then reverted to 65,000.[8] Figure 3.2 shows the growth in the number of H-1 visas issued over the last three decades, estimates of the stock of H-1 visas in the economy each year, and the changes in the H-1B visa cap.

Through the decade of the 1990s, H-1B visas became an important source of labor for the technology sector. The National Survey of College Graduates shows that 55% of foreigners working in CS fields in 2003 arrived in the US on a temporary working (H-1B) or a student type visa (F-1, J-1). Furthermore, institutional information indicates a significant increase in the number of visas awarded to computer related occupations during the 1990s. Numbers from the U.S. General Accounting Office (1992) report show that "computers, programming, and related occupations" corresponded only to 11% of the total number of H-1 visas in 1989. However, with concurrent to the Internet boom, computer scientists became a more significant fraction of individuals that received these type of working visas: according to the U.S. Immigration and Naturalization Service (2000) , the number of H-1B visas awarded to computer-related occupation in 1999 jumped to close to two-thirds of the visas, and the Department of Commerce (2000) estimated that during the late 1990s, 28% of programmer jobs went to H-1B visa holders.

While H-1B visas holders represent an important source of computer scientists, they do not represent all foreigners in the country working as computer scientists. A significant number of such foreigners are permanent immigrants, some of whom may have come either as children or as students. Other foreigners enter the US to work as computer scientists in the US on L-1B visas, which permit companies with offices both in the US and overseas to move skilled employees from overseas to the US. While we know of no data showing the fraction of computer scientists working in the US on L-1B visas, substantially fewer L-1(A&B) visas are issued than are H-1Bs.

3.2.3 The Previous Literature on the Impact of Immigrants on the High Tech Workforce in the US

Critiques of the H-1B program (e.g. Matloff (2003)) argue that firms are using cheap foreign labor to undercut and replace skilled US workers. Even the fiercest critiques of the program do not claim that employers are technically evading the law (Kirkegaard (2005)). Rather, these authors

[8]The 2000 legislation that raised the cap also excluded Universities and non-profit research facilities from it, and a 2004 change added an extra 20,000 visas for foreigners who received a masters degree in the US

argue that the requirement that firms pay visa holders the prevailing wage is close to meaningless. They claim that firms can describe positions using minimal qualifications for the job, thereby establishing a low "prevailing" wage, and then hire overqualified foreigners into the position. These authors conclude that given the excess supply of highly qualified foreigners willing to take the jobs, and given the lack of portability of the H-1B visa, workers on an H-1B visa are not in a position to search for higher wages.

One way to get a handle on the extent to which H-1B visa holders are being under-paid relative to their US counterparts is to compare foreigners on H-1B visas to those with "green cards," which are portable. Available evidence suggests that computer scientists holding green cards are paid more than observationally equivalent H-1B visa holders. Using difference-in-difference propensity score matching, Mukhopadhyay and Oxborrow (2012) find that green card holders earn 25.4 percent more than observably comparable temporary foreign workers. Additionally, based on an internet survey, Mithas and Lucas (2010), found that IT professionals with green cards earn roughly 5 percent more than observationally equivalent H-1B visa holders using log earnings regressions. Comparisons between green card and H-1B holders are far from perfect, because green cards are not randomly assigned. Many high skilled workers obtain green cards by being sponsored by their employers after they have been working on an H-1B for a number of years. It seems reasonable to assume that those being sponsored are those that both want to stay in the US and are also amongst those the employer wants to hold onto. These kind of considerations lead us to suspect that, conditional on observables, green card holders are positively selected. Given these considerations, it is somewhat surprising that the observed green card premium is not larger than it is.

While there may be no incontrovertible estimate of the productivity (conditional on earnings) advantage of foreign high skilled labor, simple economic reasons suggests this advantage must exist. US employers face both pecuniary and non-pecuniary costs associated with hiring foreigners. A small GAO survey (U.S. General Accounting Office, 2011) estimated the legal and administrative costs associated with each H-1B hire to range from 2.3 to 7.5 thousand dollars. It seems reasonable to assume that employers must expect some cost or productivity advantage when hiring foreigners. This does not mean that foreign hires are always super stars. The productivity advantage could be quite small, and could involve effort, not ability. However, without some productivity advantage, it is hard to see why employers go through the effort and expense to hire foreigners.

H-1B critics are arguing that, for the reasons discussed above, employers find hiring foreign high skilled labor an attractive alternative and that such hiring either "crowds out" natives from jobs or put downward pressure on their wages. However, as far as we know, critics of the H-1B program have not tried to estimate the magnitude of either of these effects. Recent work by economists have started to fill this void. Kerr and Lincoln (2010) and Hunt and Gauthier-Loiselle (2010) provide original empirical evidence on the link between variation in immigrant flows and innovation measured by patenting - finding evidence suggesting that the net impact of immigration is positive rather than simply substituting for native employment. Kerr and Lincoln (2010) also show that variation in immigrant flows at the local level related to changes in H-1B flows do not appear to adversely impact native employment and have a small, statistically insignificant effect on their wages.

A potential issue with Kerr and Lincoln's analysis is that the observed, reduced-form outcomes may capture concurrent changes in area specific demand for computer scientists. Kerr and Lincoln fully understand this endogeneity issue. To circumvent the problem, they construct a variable that interacts an estimate for the total number of individuals working on H-1B visas in a city with local area dependencies on H-1Bs. Their hope is that the variation in this variable is driven largely by changes in the cap on new H-1B visas that occurred over the last 20 years. That said, it is unclear the extent to which the variation Kerr and Lincoln use is being driven by variation in the visa cap. Because of the dot com bubble bust in 2000 and 2001, the variation in the H-1B cap is only loosely related to actual number of H-1Bs issued. In addition, it is hard to imagine that the cap was exogenous to the demand for IT workers. Finally, if because of local agglomeration effects, the IT boom was concentrated in areas of the country that were already IT intensive (such as Silicon Valley), then the measure of local dependency would be endogenous.

In the context of an economic model, it is difficult to generate a situation in which there is little crowd out unless labor demand is very elastic. While there are models of the labor market which could rationalize such large elasticities,[9] this paper proposes an alternative interpretation

[9]If computer scientists have large effects on firm productivity, then demand curves for them would be very elastic.

to Kerr and Lincoln's results, even when the labor demand is not close to perfectly elastic. If employers face costs to hire immigrant labor and are bound to pay the going wage, firms might disproportionately hire immigrants only when the demand for workers is increasing. In this case, immigrants would not replace incumbent workers or depress wages, but rather have a negative impact on the growth of wages and employment for natives. Under these circumstances, one might very well see a positive association between an increase in the utilization of foreign computer scientists and the increased utilization of their US counterparts, even though the availability of skilled foreigners is putting downward pressure on the growth in earnings and employment of native computer scientists.

3.3 A Dynamic Model of Supply and Demand of Computer Scientists

To gauge the impact that the availability of foreign high skilled labor has had on US workers, we construct a simple model of the labor market for computer scientists. While our model is quite stylized, we intend to capture the most salient features of the market.

In our model there are three potential sources for computers scientists. First, there are those who earn computer science bachelor's degrees from US institutions. These individuals must complete college before they are ready to work. Second, there are US residents working in other occupations who can switch into computer science, but must pay costs to switch occupations. Third, there are foreigners who are being recruited on temporary work visas.[10] There is also the group who immigrated with their parents as children, but these individuals are typically either citizens or green card holders and we assume employers do not distinguish between these individuals and the US born. We also ignore the fact that some immigrants are coming in on permanent visas. As the GAO and Department of Commerce reports cited earlier suggest, at least in the 1990s, the majority of foreigners working as computer scientists within the US who have finished their undergraduate degrees abroad, arrived on temporary work visas. In addition, the data we will use does not allow

Alternatively, one could imagine that, absent the foreign computer scientists, production would shift overseas either because of domestic firms outsourcing production or because of Heckscher-Ohlin effects.

[10] Here we are aggregating foreign students getting degrees in the US with their domestic counterparts. During the 1990s, foreigners represented a small (10%) share of new CS graduates each year (IPEDS completion survey).

us to distinguish visa types.

In terms of the demand side of the model, we assume that firms observe the technological progress level and make decisions about whether to hire foreigners or domestic workers. We assume that foreigners are somewhat more productive than US workers but are paid the same wage due to institutional restrictions. Alternatively, we could have equally well assumed employers experience a cost advantage associated with hiring foreigners. Furthermore, firms face increasing costs for recruiting foreigners, making it non-optimal for firms to only hire foreign workers.

3.3.1 Labor Supply of American Computer Scientists

We model U.S computer scientists as making two types of decisions along their career in order to maximize the expected present value of their life time utility. At age 20, individuals in college choose the field of study that influences their initial occupation after graduation, and from age 22 to 65, workers choose between working as a computer scientist or in another occupation. Individuals have rational, forward looking behavior and make studying and working decisions based on the information available at each period.

3.3.1.1 Studying decision

We assume that students make their major decisions when they are juniors in college. At age 20, an individual i draws idiosyncratic taste shocks for studying computer science or another field: η_i^c and η_i^o,respectively. This student also has expectations about the prospects of starting a career in each occupation after graduation (age 22), which have a values V_{22}^c and V_{22}^o respectively. With this information, an individual chooses between pursuing computer sciences or a different choice of major at the undergraduate level.[11]

We model the utility of a student as a linear function of the taste shocks and career prospects in each sector. There is also a taste attractiveness parameter α_o for studying a different field from computer science and individuals discount their future with an annual discount factor β. With these

[11]Essentially, we are assuming that students decide their major after the end of their second year in school. This presumes that the relative pool of potential applicants would have sufficient background to potentially major in computer science. A four year time horizon is more standard. We experimented with such a horizon and doing so made little qualitative difference to our conclusions.

assumptions, the field of study decision is represented by:

$$max\{\beta^2\mathbb{E}_t V_{22}^c + \eta_i^c, \beta^2\mathbb{E}_t V_{22}^o + \alpha_o + \eta_i^o\}$$

We assume that η_i^c and η_i^o are independently and identically distributed and for $s = \{c, o\}$, can be defined as $\eta_i^s = \sigma_0 v_i^s$, where σ_0 is a scale parameter and v_i^s is distributed as a standard Type I Extreme Value distribution. This distributional assumption is common to dynamic discrete choice models (Rust (1987), Kline (2008)) and it is convenient because it allows the decisions of agents to be smoothed out, a desired property that will be used in the characterization of the equilibrium of the model.

Given the distributional assumption of idiosyncratic taste shocks, it follows that the probability of a worker *graduating* with a computer science degree can be written in logistic form:

$$p_t^c = [1 + exp(-(\beta^2\mathbb{E}_{t-2}[V_{22}^c - V_{22}^o] - \alpha_o)/\sigma_0)]^{-1}$$

Note that the important parameter for how studying choices of workers are sensitive to different career prospects is the standard deviation of taste shocks. Small values of σ_0 imply that small changes in career prospects can produce big variations in the number of students graduating with a computer science degree.

The next step to characterize the supply of young computer scientists is to map the graduating probability described above to employment. Defining M_t^a as the exogenous number of college graduates with age a in time period t,[12] the number of recent graduates with a computer science degree in year t is represented by $C_t = p_t^c M_t^{22}$.

3.3.1.2 Working Decision

The field of study determines if an individual enters the labor market as either a computer scientist or with a different occupation. However, individuals can choose to switch occupations along their careers. Specifically, at the beginning of each period, individuals between ages 22 and 65 choose to

[12] We are implicitly assuming that anyone who majors in computer science would have completed college even had they not majored in computer science and that computer science majors are infra marginal college finishers. A similar assumption was made by Ryoo and Rosen (2004) in their work on Engineers.

work in CS or another type of job in order to maximize the expected present value of their lifetime utility.

A feature of the model is that switching occupations is costly for the worker. A justification for this assumption is that workers have occupational-specific human capital that cannot be transferred (Kambourov and Manovskii (2009b)). We assume the cost to switch occupations is a quadratic function of a worker's age. Note that this assumption implies that it becomes increasingly harder for workers to switch occupations as they get older. Additionally, there is no general human capital accumulation and wages do not vary with the age of a worker.[13]

Finally, we assume that workers have linear utility from wages, taste shocks and career prospects. Furthermore, wages must be totally consumed in that same year and workers cannot save or borrow. The Bellman equations of worker i at age a between 22 and 64 at time t if he starts the period as a computer scientist or other occupation are respectively:

$$V^c_{t,a} = max\{w^c_t + \beta\mathbb{E}_t V^c_{t+1,a+1} + \varepsilon^c_{it}, w^o_t - c(a) + \beta\mathbb{E}_t V^o_{t+1,a+1} + \varepsilon^o_{it} + \alpha_1\}$$

$$V^o_{t,a} = max\{w^c_t - c(a) + \beta\mathbb{E}_t V^c_{t+1,a+1} + \varepsilon^c_{it}, w^o_t + \beta\mathbb{E}_t V^o_{t+1,a+1} + \varepsilon^o_{it} + \alpha_1\}$$

where $c(a) = \lambda_0 + \lambda_1 a + \lambda_2 a^2$, is the monetary cost of switching occupation for an age a worker, and α_1 is the taste attractiveness parameter for not working as a computer scientist. For simplicity, we assume that the current wage in the other occupation w^o_t is exogenous and perfectly anticipated by the workers.[14] In the model, all workers retire at age 65 and their retirement benefits do not depend on their career choices. As a consequence, workers at age 65 face the same decision problem but without consideration for the future.

As in the college-major decision problem, idiosyncratic taste shocks play an important role in working decisions of an individual. Once more, we will assume that taste shocks are independently[15] and identically distributed and for $s = \{c, o\}$ can be defined as $\varepsilon^s_{it} = \sigma_1 v^s_{it}$ where σ_1 is a

[13]The implications of the model will still hold if there is general human capital accumulation and individuals expect similar wage growth profiles working as computer scientists or in the alternative occupation.

[14]As a matter of fact, in the simulations of the paper we will set $w^o_t = 1$ and measure wages of computer scientists as an occupational premium.

[15]In the working decision problem, the independence assumption might be less plausible because taste shocks could be serially correlated. However, identifying parameters of the model with serially correlated errors is infeasible without longitudinal data (Kline (2008)).

scale parameter and v_i^s is distributed as a standard Type I Extreme Value distribution.

Defining $p_{t,a}^{sS}$ as the probability that a worker at age a between 22 and 64 moves from occupation s to occupation S, it follows from the error distribution assumption that the migration probabilities can be represented as:

$$p_{t,a}^{oc} = [1 + exp(-(w_t^c - w_t^o - c(a) - \alpha_1 + \beta \mathbb{E}_t[V_{t+1,a+1}^c - V_{t+1,a+1}^o])/\sigma_1)]^{-1}$$

$$p_{t,a}^{co} = [1 + exp(-(w_t^o - w_t^c - c(a) + \alpha_1 + \beta \mathbb{E}_t[V_{t+1,a+1}^o - V_{t+1,a+1}^c])/\sigma_1)]^{-1}$$

and the migration probabilities of workers at age 65 are the same without discounting future career prospects. Note that the switching probabilities depend upon both the current wage differential and expected future career prospects at each occupation. The standard deviation of the taste shocks, the sector attractiveness constant and the cost of switching occupations will effect the extent to which changes in relative career prospects affect the movement of US residents across fields.

A feature of dynamic models with forward looking individuals is that working decisions depend upon the equilibrium distribution of career prospects. As in the dynamic choice literature with extreme value errors (Rust (1987) and Kline (2008)), we use the properties of the idiosyncratic taste shocks distribution to simplify the expressions for the expected values of career prospects. As a result, the expected value function for an individual at age a between 22 and 64 working as a computer scientists or in another occupation are respectively:

$$\begin{gathered}
\mathbb{E}_t V_{t+1,a+1}^c = \\
\sigma_1 \mathbb{E}_t[\gamma + ln\{exp((w_{t+1}^c + \beta \mathbb{E}_{t+1} V_{t+2,a+2}^c)/\sigma_1) + exp((w_{t+1}^o - c(a) + \alpha_1 + \beta \mathbb{E}_{t+1} V_{t+2,a+2}^o)/\sigma_1)\}] \\
\\
\mathbb{E}_t V_{t+1,a+1}^o = \\
\sigma_1 \mathbb{E}_t[\gamma + ln\{exp((w_{t+1}^o + \alpha_1 + \beta \mathbb{E}_{t+1} V_{t+2,a+2}^o)/\sigma_1) + exp((w_{t+1}^s - c(a) + \beta \mathbb{E}_{t+1} V_{t+2,a+2}^c)/\sigma_1)\}]
\end{gathered} \tag{1}$$

where gamma $\gamma \cong 0.577$ is the Euler's constant and the expectations are taken with respect to future taste shocks. Workers at age 65 face the same expected values but do not discount the future.

Now we turn to transforming migration probabilities to employment. The first step is to determine the CS supply of recent college graduates. After leaving college, individuals can start their careers in the occupation corresponding to their field of study with no cost. However, we also allow workers at age 22 to pay the switching costs and get their first job in an occupation different from their field of study. As a consequence, the number of computer scientists at age 22 is a function of the number of recent graduates with a computer science degree and the migration probabilities:

$$L_t^{22} = (1 - p_{t,22}^{co})C_t + p_{t,22}^{oc}[M_t^{22} - C_t]$$

where M_t^{22} is the number of recent college graduates, C_t is the number of recent graduates with a computer science degree, and $M_t^{22} - C_t$ is the number of college graduates with any other degree.

In the same way, the supply of computer scientists at age a from 23-65 is a function of past employment in each occupation and the migration probabilities:

$$L_t^a = (1 - p_{t,a}^{co})L_{t-1}^{a-1} + p_{t,a}^{oc}[M_{t-1}^{a-1} - L_{t-1}^{a-1}]$$

where M_t^a is the exogenous total number of workers in the economy at age a in time period t. $M_t^a - L_t^a$ is the number of workers at age a working in the residual sector. For simplicity, we assume that the number of workers in the economy at age M_t^a is exogenous and constant over time.[16]

The aggregate domestic labor supply of computer scientists is the sum of labor supply at all ages:

$$L_t = \sum_{a=22}^{a=65} L_t^a \text{ (2)}$$

Note that the labor supply of computer scientists depends on past employment, new college graduates with a computer science degree and on wages through the migration probabilities.

3.3.2 Labor Supply of Foreign Computer Scientists

An important characteristic of our model is that firms can recruit foreigners to work as computer scientists. As it will become clear throughout the section, this possibility has implications on how

[16]In the simulation of the paper we set M_t^a to be constant for all ages and $\sum_{a=22}^{a=65} M_t^a = 100$. We measure employment of computer scientists as percentage points of the employed population of interest.

the market for CS workers responds to technological shocks, such as Internet innovation, in terms of enrollment decisions, wages and employment.

We model foreign computer scientists as having a perfectly elastic labor supply. The wage that a computer scientist could obtain in India, for example, is so much lower than it is in the US that the wage premium creates a large queue of individuals ready to take jobs in the US (Clemens (2013) provides direct evidence on this point).[17] Additionally, we assume that foreigners cannot switch their occupation once hired to work as computer scientists and they continue to work in the US until their visa expires.[18]

A simplified way to model the framework describe above is to define R_t as the number of foreigners recruited as CS in period t. Next, we assume that all CS foreigners stay in the US for 6 years, that is the maximum length of a H-1B visa contract.[19] In this framework, the number of foreigners currently working as CS in the US is defined as the sum of current and the recruitment in the past 5 years:

$$F_t = \sum_{j=0}^{5} R_{t-j} \quad (3)$$

3.3.3 Labor Demand for Computer Scientists

We model the labor demand as resulting from the decisions made by a standard representative firm in a perfectly competitive framework. In the model, firms observe both the wage and technological progress levels and choose US and foreign employment in order to maximize their intertemporal profits. While firms do not assume that their US employees will necessarily stay with them from one period to the next, given the institutional setting, firms do assume that foreign workers will continue with the firm until the workers' visa expires six years after he or she is hired.

We assume there is only one type of firm that hires computer scientists. CS labor is the only

[17]As it will become clear later, the reason why in our model foreigners do not swamp the U.S. labor markets is because firms must pay, in addition to prevailing wages, increasing recruitment costs to employ foreigners.

[18]In fact, during the period we are studying roughly half of those on H-1B visas eventually became permanent residences. In our online appendix, we present a modification of the model that allows a constant fraction of H-1B visa holders to become permanent residents. Our results are consistent across modeling specifications.

[19]The initial duration of the H-1B contract is 3 years, but it is extendable for an additional 3 years. Extensions do not count toward the H-1B cap, and are generally granted. As it will become clear in the labor demand side, in our model firms have incentive to keep foreigners for the maximum length of their contract.

input used in the production function and we ignore the firm's decision about capital or other types of labor adjustments.[20] We further assume that computer scientists at different ages are perfect substitutes in the production function. As a consequence, firms do not distinguish workers by age when making their hiring decision, precluding the kind of issues addressed by Kerr et al. (2013).[21] In addition, we assume that foreigners and US workers are close substitutes in the production function, but foreigners have higher marginal productivity than US workers.

A restriction we impose in the model is that all computer scientists in the market are paid the same wage independently of their age or citizenship. This assumption is in accordance with the H-1B visa regulation that requires that wages paid to foreigners must be at least the prevailing wage rate for the occupational classification in their area of employment. Finally, there are no adjustment costs for American workers but firms incur extra costs to recruit foreigners.[22] This expenditure is justified by the fees and expenses directly related to the visa application process, and also the extra cost that a firm typically has for searching for workers overseas.

As it will become clear throughout the section, this framework implies that firms face a trade-off when making the decision of hiring foreigners. On one hand, foreigners have a higher marginal productivity than US workers and are paid the same wage. As a consequence, firms are willing to substitute foreign workers for their US workers. On the other hand, there are extra recruitment costs to bring foreigners to the US. This restriction implies that firms never completely substitute foreign for US workers.

3.3.3.1 Firm's Decision

The forward looking firm makes decisions about the recruitment of US and foreign workers in order to maximize intertemporal profits, as represented by the Bellman equation:[23]

[20]The assumption that labor adjustment decisions are independent of capital is standard in the dynamic labor demand literature when data on capital stock is not available (Hamermesh (1989)). Including capital in the production function generally does not qualitatively change the results (Kline (2008)).

[21]While we suspect is would make sense to allow workers of different ages to be imperfect substitutes in production for each other, CPS sample sized are too small to support this kind of analysis.

[22]In our online appendix we set-up and calibrate a model where the quadratic cost term for hiring foreigners also applies to Americans. Our results are not sensitive to this modeling change.

[23]For simplicity, we assume that firms and individuals have the same annual discount factor β. For expositional purposes, we now omit the the superscript c for wages and employment of computer scientists.

$$\pi_t = max_{L_t,R_t} A_t Y(L_t + \theta F_t) - w_t(L_t + F_t) - C_R(R_t) + \beta \mathbb{E}_\mathrm{t}[\pi_{t+1}]$$

subject to foreign labor supply:

$$F_t = \sum_{j=0}^{5} R_{t-j}$$

where $A_t Y(.)$ is the production function, θ is a constant greater than 1 that represents marginal productivity differences between foreigners and US workers, and $C_R(.)$ is the recruitment cost function of foreigners.

We represent the production function as Cobb-Douglas, such that $Y(L_t + \theta F_t) = (L_t + \theta F_t)^\gamma$, for some γ between zero and one, implying a downward sloping labor demand curve for computer scientists. This set-up can be made consistent with the Romer (1986) model of knowledge accumulation as a by-product of capital accumulation; or the Arrow (1962) learning-by-doing model, where we allow increases in employment to lead to increases in productivity. To see this, we can reformulate the production function to be $Y_t = [B_t(L_t + \theta F_t)]^\delta$. If we let the technology parameter exhibit learning-by-doing, then $B_t = \psi_t(L_t + \theta F_t)^\alpha$, giving us a production function of the form $Y_t = \psi_t^\delta (L_t + \theta F_t)^{\delta\alpha}$. If we define, $A_t = \psi_t^\delta$and $\gamma = \alpha\delta$, then we recover the simple Cobb-Douglas production function: $A_t(L_t + \theta F_t)^\gamma$. The parameter, γ, should then be thought of as a reduced-form parameter that captures not just the effective labor share in output, but also the productivity gains from hiring more effective workers. As long as γ lies between 0 and 1, this parametrization guarantees a decreasing marginal return to labor and thus an interior solution for the employment decision of the firm. Furthermore, the parameter γ has a direct mapping to the long-run elasticity of labor demand with respect to effective labor ($L_e = L + \theta F$):

$$\epsilon_{L_e,w} = \frac{1}{1-\gamma}$$

Additionally, we assume that recruitment costs of foreigners include both linear and quadratic components $C_R(R_t) = c_1 R_t + c_2 R_t^2$. The linear term in the foreign recruitment cost represents expenditures that are required for hiring each foreign worker, such as application fees. The quadratic term has been widely used in dynamic labor demand literature (Sargent (1978) and Shapiro (1986)). As will become clear from the first order condition of the firm, convex hiring costs prevents firms from completely substituting foreigners for domestic workers, by increasing marginal recruitment costs of foreigners.[24]

As in a typical dynamic labor demand problem the solution to the firm's decision can be characterized by both the first order and envelope conditions with respect to the employment level. The first order condition of the firm's maximization problem with respect to US employment is represented by the following equation:

$$A_t\gamma(L_t+\theta F_t)^{\gamma-1} = w_t \text{ (4)}$$

Note that because there is no adjustment costs for US workers, the first order condition with respect to US employment is the same as in a static maximization problem. It is simply characterized by firms equalizing the marginal product of US workers to their wage level.

In addition to choosing US worker employment, the firm also decides the number of foreign workers recruited at each period. The first order condition of the firm's problem with respect to R_t is given by:

$$\theta A_t\gamma(L_t+\theta F_t)^{\gamma-1} - w_t - c_1 - 2c_2R_t + \sum_{j=1}^{5}\beta^j\mathbb{E}_{\mathrm{t}}[\tfrac{\partial\pi_{t+j}}{\partial R_t}] = 0$$

where $\frac{\partial\pi_{t+j}}{\partial R_t}$ is defined as how profits in $t+j$ are affected by changes in the recruitment in t. Finally, we use envelope condition to derive the shadow price of past foreign recruitment on current profits, such that:

$$\tfrac{\partial\pi_t}{\partial R_{t-j}} = \theta A_t\gamma(L_t+\theta F_t)^{\gamma-1} - w_t \text{ for } j = 1, ..., 5$$

Rearranging the first order and envelope conditions of foreigner recruitment leads us to the useful alternative representation to the demand for foreign workers:

$$\sum_{j=o}^{5}\beta^j\mathbb{E}_{\mathrm{t}}[\theta A_{t+j}\gamma(L_{t+j}+\theta F_{t+j})^{\gamma-1} - w_{t+j}] = c_1 + 2c_2R_t \text{ (5)}$$

[24]Our formulation implies the foreign share of new hires will rise as demand increases. There are alternative models that would imply something similar. For example, if firms had some local monopsony power, and if foreign labor were supplied elastically, firms would accommodate demand increases by shifting recruitment toward foreign labor so as to avoid paying increased wages associated with the increased hiring of US trained labor.

Equation (5) shows the trade-off faced by firms when hiring foreigners. The left hand side can be interpreted as the present value of the expected marginal benefit of recruiting a foreigner, defined as the difference between the marginal productivity of a foreigner and wage level during the 6 years duration of his contract. Note that firms benefit from hiring foreigners because they are more productive than US workers by a constant θ but are paid the same wage. The right hand side represents the marginal cost of recruiting a foreigner. Since the marginal cost of recruiting a foreigner is increasing with R_t, firms will never completely substitute foreigners for US workers in the model.

3.3.4 Equilibrium

A dynamic general equilibrium can be characterized by the system of equations that represent those choice functions and the stochastic process of technological progress A_t. In particular, equation (1) characterizes the expectations of workers with respect to future career prospects, equations (2) and (3) are the dynamic labor supply of American and foreigner computer scientists respectively, and equations (4) and (5) describe the dynamic labor demand for American and foreign CS.

The last piece to characterize the equilibrium of the model is to define a stochastic process of technological progress. Note that A_t is the only source of exogenous variation to the system. We choose to specify A_t as a close to random walk process,[25] such that:

$$A_t = 0.999A_{t-1} + 0.001\bar{A} + \xi_t \text{ (6)}$$

where $\bar{A}$ is the steady state level of progress, and ξ_t is the i.i.d. random idiosyncratic productivity

[25] We model the technology progress as a close to random walk since we will interpret the Internet boom as a series of very persistent technological shocks that hit the information technology sector during the late 1990s. We also interpret the 2000 to 2004 to be a dot com bust. We found little evidence that workers, students or employers expected the increase in the demand for computer scientists during the 1990s to be temporary (and subject to a post-Y2K bug slump). First, the BLS projected a steady increase in CS employment after the year 2000, and claimed that it expected the top two fastest growing occupations to be computer scientists, and computer engineers respectively. Furthermore, there is a substantial increase in CS degrees started during the dot-com boom, indicating that students perceived the demand for computer scientists to be increasing permanently during the period. We therefore believe that a more realistic assumption is that agents perceived the increase in demand during the late 1990s to be permanent - and that the World Wide Web generated opportunities for new businesses that demanded computer scientists. However, at some period in the beginning of the year 2000, presumably for a variety of reasons, the boom turned around and NASDAQ crashed.

shock with mean zero that is assumed to be independent of other variables of the model.[26]

The equilibrium of the model can be expressed by a mapping from the state variables: $s = \{\mathrm{C_t}, L_{t-1}^{22}, ..., L_{t-1}^{64}, R_{t-1}, ..., R_{t-5}, A_{t-1}\}$ and exogenous productivity shock ξ_t to the values of L_t, w_t, R_t, and $\mathbf{Vt}$, the vector of career prospects at different occupations for different ages, that satisfies the system of equations (1) to (6). We solve the system by numerically simulating the model in Dynare (a widely used software) via perturbation methods (Juillard (1996)). The policy functions are calculated using a second order polynomial approximation to the decision rules implied by the equations of the model Collard and Juillard (2001a,b).

3.4 Calibration and Simulation

3.4.1 Identification and Calibration Method

There are twelve parameters in the model $\{\sigma_0, \alpha_0, \sigma_1, \alpha_1, \lambda_0, \lambda_1, \lambda_2, \beta, \gamma, \theta, c_{R1}, c_{R2}\}$. We set the foreign worker productivity[27] parameter $\theta = 1.12$ based on estimations from the 2003 National Survey of College Graduates data.[28] This value of the wage premium earned by foreign green card holders is broadly consistent with other estimates in the literature (Mithas and Lucas (2010), Mukhopadhyay and Oxborrow (2012)). Furthermore, we set the annual discount rate of workers and firms $\beta = 0.9$. Our results are not sensitive to plausible variations of this parameter.

In our modeling we are treating the wage, employment and enrollment shifts as a response to an exogenous shift in the demand for computer scientists due to the technological developments that occurred during the period of analysis. We use this demand shift to identify the enrollment and labor supply response of natives, and the parameters affecting the hiring decision of foreigners: $\{\sigma_0, \alpha_0, \lambda_0, \lambda_1, \lambda_2, \sigma_1, \alpha_1, c_{R1}, c_{R2}\}$. At the same time, demand shifts will not identify the slope of

[26]Note that both workers and firms are risk neutral in our model. For this reason, the certainty equivalence property holds and the solution of the model does not depend on higher moments of the idiosyncratic productivity shock.

[27]In an Online Appendix we re-do all our results for different values of this parameter, and find that our results are not sensitive to the choice of this parameter.

[28]Specifically, we estimate the wage premium for foreign born computer scientists who are naturalized or permanent residents, compared to US born CS workers. This estimation comes from a logarithmic of annual earnings regression controlling for gender and a cubic age polynomial. We interpret this wage premium as the average marginal productivity difference between foreign and US computer scientists.

the labor demand curve. As a result, we present the results of the paper using different assumptions about the values of γ.

To calibrate $\{\sigma_0, \alpha_0, \lambda_0, \lambda_1, \lambda_2, \sigma_1, \alpha_1, c_{R1}, c_{R2}\}$, we use observations of US and foreign employment, wages, and enrollment[29] between 1994 and 2004. We define other STEM occupations as the career alternative to CS jobs. The data we are using on employment and earnings is derived from the March Current Population Survey. This survey contains no indication as to the visa status of the foreign born. To approximate the population of interest, we identify the foreign born who immigrated to the US after they turned 18 as our foreign workers. We also normalize employment variables to use units of American STEM workers, and wages to use units of wages[30] of other STEM jobs, and thus define our key data series as:[31]

1. $L_t = \frac{\text{US computer scientists}}{\text{US workers with STEM occupations}}$

2. $F_t = \frac{\text{Foreign computer scientists}}{\text{US workers with STEM occupations}}$

3. $w_t = \frac{\text{Median weekly wages for computer scientists}}{\text{Median weekly wages for other STEM jobs}}$

4. $p^c_{t+2} = \frac{\text{US computer science Bachelor's degrees awarded (lagged 2 years)}}{\text{US STEM Bachelor's degrees awarded (lagged 2 years)}}$

5. $s_t^{a_1,a_2} = \frac{\text{US computer scientists with age between } a_1 \text{and } a_2}{\text{US computer scientists}}$

   For $a_1$and a_2 defined as the age ranges $\{$22 to 34; 35 to 44; 45 to 65$\}$.

Conditional on γ and θ and observations of $\{w_t, L_t, F_t\}$ we are able to recover values of A_t implied by our model during the period of 1994 to 2004:

$$A_t = \frac{w_t}{\gamma(L_t+\theta F_t)^{\gamma-1}}$$

[29]Uses data from 1996 to 2006, representing enrollment decisions from 1994 to 2004. See the Online Appendix for more details.

[30]We exclude imputed values of wages, and multiply top-coded values by a factor of 1.4. Bollinger and Hirsch (2007) show that not excluding imputations can lead to biased results. Whereas the top-coding adjustment is standard in the literature (Lemieux (2006)). See the Online Appendix for more details.

[31]See the data appendix online for more information on occupational classifications. We smooth the raw data as follows: $X_{t,smooth} = \frac{1}{3}(X_{t-1,raw} + X_{t,raw} + X_{t+1,raw})$, except for the American and foreigner employment data in 1994, which citizenship information is unavailable prior to 1994 for which we use: $X_{1994,smooth} = \frac{2}{3}X_{t,raw} + \frac{1}{3}X_{t+1,raw})$.

We assume that the economy is in steady state in 1994, such that $\bar{A} = A_{1994}$, and that it is hit by the series of shocks. In terms of expectations, we assume that both firms and individuals are surprised by changes in A_t.[32] Note that following equation (6), firms and workers have essentially static expectations about future technology progress, such that $\mathbb{E}_t[A_{t+j}] \cong A_t$ for any j.

The remaining parameters $\{\sigma_0, \alpha_0, \lambda_0, \lambda_1, \lambda_2, \sigma_1, \alpha_1, c_{R1}, c_{R2}\}$ are calibrated such that the model matches the observations of L_t, F_t, w_t, in two periods of time: 1994 and 2004, and the changes in the age structure $s_t^{a_1,a_2}$ in 2004.[33] We use a Nelder-Mead simplex method to find parameter values which yield solutions to the model under these data restrictions.[34] The intuition for the identification of the parameters comes straight from the data. For the given series of exogenous technological shocks and wages, variations of enrollment between 1994 and 2004 identify the parameters σ_0 and α_0, changes in native employment identify the parameters σ_1 and α_1, variations in foreign employment identify the recruitment cost parameters c_{R1} and c_{R2}, and changes in the age structure of computer scientists identify the quadratic costs of switching occupations: parameters λ_0, λ_1, and λ_2.

3.4.2 Calibration results

We use the procedure described above to calibrate the model using three different values of γ: {0.25, 0.5, 0.75}.[35] We present the calibrated parameters for these different values of γ in Table

[32]We also considered the alternative assumption that all agents fully or partially anticipated the future path of technological process. This assumption yields time paths for wages and employment that are quite similar to the ones we observe under our static expectations assumption. In contrast, with this alternative assumption, enrollment jumps almost immediately, which is inconsistent with the time path of enrollment we observe. At the same time, our counterfactual simulations presented later with the alternative anticipation assumption are similar to the ones we present with static expectations. Presumably a model that allowed expectations to evolve would be more realistic. However, given the robustness of our central results to the static versus foresight assumption, we did not explore such an alternative.

[33]The decision to match changes in the age structure of CS rather then levels is to assure that our calibrated model reflects movements that occurred in the market for CS during the period rather than the age structure of the entire population.

[34]Note that we have a perfectly identified system: we find the values of 9 independent parameters and 2 implied values of A_t such that the model matches 11 data observations: L_t, F_t, w_t and p^c_{t-2} in two years and the observation of changes in $s_t^{22,34}$, $s_t^{35,44}$,and $s_t^{45,65}$in 2004.

[35]γ in the 0.25 to 0.75 range imply labor demand elasticities between -1.33 and -4.0. Ryoo and Rosen (2004a), estimate demand elasticities for engineers that lie between -1.2 and -2.2, while Borjas (2009), studying the effect the immigration of foreign born PhD scientists on the wages of competing workers, estimates demand elasticities of approximately -3.0. This do suggest that we have varied γ through a sensible range.

3.2 and a comparison of the data with the model's output in Figures 3.3 - 3.4. We consider the demand elasticities derived from our γ's to span a reasonable range of plausible values of this parameter, which as we describe in Section 3.4.4, is challenging to identify.

The calibrated model allows us to calculate several additional economically meaningful statistics, which we also include in the bottom segment of Table 3.2. We calculate the long-run occupation and enrollment elasticities with respect to wages, by replacing the demand side of the model with an exogenous wage, which we set to be permanently 1% higher than its 1994 value, and in each case, we allow the supply side to come to a new equilibrium based on the calibrated parameters.We similarly calculate the short-run occupation and enrollment elasticities, but instead of allowing the supply side to come to a new steady-state, we calculate the elasticities based off of changes in occupation and enrollment after 1 year.

In the bottom section of Table 3.2, we show how each of these long-run elasticities varies with γ. The long-run occupational labor supply elasticity for Americans is around 5.4. The enrollment in CS is even more elastic, with a long-run elasticity that lies around 11.6.[36] This result reflects the large enrollment response we witness in the data. The short-run occupation elasticity is much lower than the corresponding long-run elasticity. We expect this result, due to the supply frictions and lags in our model. In contrast, the short-run and long-run enrollment elasticities are almost exactly the same. Pre-enrollment students respond immediately to a wage shock. A fuller model which includes capacity constraints on the supply side of the higher education market, would work to slow such adjustments. Finally, the average cost of recruiting a foreign worker is about 0.53 times the average annual earnings of a non-CS STEM job.

In Figures 3.3-3.4, we report both the path predicted by our calibrated model (Full model) and the path observed in the data (Smooth data) during 1994-2009. Note that by the construction of our calibration procedure, the full model fits the data perfectly in 1994 and 2004. We use the transition period between 1995 to 2003 to evaluate how well the model fits the data, and the years 2005-2009 for out of sample prediction. These years include observed changes to relevant immigration laws, and potentially unobserved structural changes which would map to changes in our parameters, so

[36]Ryoo and Rosen estimate substantially smaller enrollment elasticities of between 2.5 and 4.5, but are modeling the decision to enroll in a broader field than we are.

our model has trouble fitting the data in this period for some series. Figure 3.3 shows that for different γ's, the model is a fairly close fit for CS wages and American employment during the evaluation period, although CS wages in the model grow faster at first and American employment in CS grows more slowly in the model than the data. The fit of these two series is still relatively good in the out of sample prediction period, with wages slightly higher and American employment slightly lower in the model compared to the data.

Figure 3.4 shows that the enrollment output of the model is particularly sensitive to the choice of γ, where lower values somewhat under-predict the enrollment boom surrounding 2001. At odds with the predictions of our model, enrollment does not increase starting in 2006. Given the rising wages of computer scientists at the time, this pattern seems a bit surprising and we confess to not having a good understanding as to why enrollments do not seem to be responding to market signals. The figure also shows that foreign employment grows more slowly at first in the model than the data. In the out of sample period, foreign employment shrinks in the model instead of growing slightly, as in the data. This could be because our model assumes that after a 6-year period, foreigners return to their home country. In the Online Appendix, we calibrate a model that allows a certain fraction of H-1B workers to become permanent residents. This extension of the model does a better job of fitting the share of foreign employment in the last few years (and overall does a good job of fitting the different calibrated series).

3.4.3 Simulation of Fixed Foreign Worker Population Counterfactual

We use our calibrated model to simulate a counterfactual Internet boom from 1994-2004, as if firms had restrictions on the number of foreigners that they can hire. The exercise consists of hitting the calibrated model with the same technological shocks we derived before but imposing that firms cannot increase F_t above its 1994 level. The results of this simulation are also presented in Figures 3.3-3.4 (Restricted Model). There we can compare the counterfactual for different values of γ with the smoothed data.

Overall, our calibrated model implies an increase in the demand for domestic workers when firms cannot increase foreign employment above its 1994 level. As a result, we observe higher wages, US employment and computer science enrollment in the counterfactual economy. We simulate significant differences in the labor market for computer scientists during the Internet boom if firms had restrictions on the number of foreigners they could hire. While the data shows that the relative wages for CS workers increased by 3.2% between 1994 and 2004, in the simulated economy wages would have increased between 5.9% to 6.9% (decreasing with γ) during the same period. In terms of employment, we observe an increase of 41% of total CS employment during the Internet boom, while in the economy where we restrict foreign workers we find an increase of only 29.1% to 36.1% (decreasing with γ) during the same period. This change in employment results from the more inelastic labor supply curve that firms face when extra foreigners are not available.

In Table 3.3 we compare the 2004 levels of the variables of interest between the data and the simulated economy where firms could not increase foreign employment above its 1994 levels. We find that in 2004, CS workers wages would be 2.8% to 3.8% higher if firms had restrictions in the number of foreigners they could hire. Furthermore,the number of Americans working in the CS sector would be 7.0% to 13.6% higher in 2004, but the total employment level would be lower by 3.8% to 9.0%. Finally we find a significant difference in the number of students enrolling in computer science in the simulated counterfactual economy. Relative to other STEM fields, enrollment in CS would be 19.9% - 25.5% higher in 2004 if firms could not increase foreign employment during the Internet boom. These numbers reflect the fact that, according to our calibrations, students' major choices are very sensitive to changes in wages.

To sum up, even when assuming a very elastic labor demand curve (high γ values) we find significant effects of foreign recruitment on wages and employment of domestic CS workers during the Internet boom. Additionally, firms would not replace all foreigners with domestic workers during this period if they were restricted to keeping the same foreign employment level of 1994, implying that industry output would be reduced.

3.4.4 Identification of Labor Demand

As shown previously, the labor market outcomes of the counterfactual simulations holding F_t fixed can vary with values of γ. In particular, we observe that when using a more elastic labor demand (higher γ), our simulated counterfactual economy (where we restrict foreigner workers) from section 4.3 has smaller increases in wages and US employment. The natural question is which, if any, of the 3 different $\gamma's$ yields results that are closest to what we would observe if firms had not been able to recruit foreigners during the Internet boom?

In a closed, constant returns to scale economy, the elasticity of demand for computer scientists would depend on both the substitutability between consumption of goods produced in sectors of the economy intensive in computer scientists and other goods, and on the substitutability between production of computers scientists and other factors of production. Given the fact that the share of computer scientists working in any one sector is not large,[37] the demand elasticity will be determined largely by the elasticity of substitution between computer scientists and other factors of production. In the relatively small window of time we are talking about, it is hard to believe these elasticities are that large.

There are two factors that mitigate this basic conclusion. First, to the extent that computer scientists contribute to innovation in the sectors of the economy intensive in computer scientist labor, the derived elasticity of demand for computer scientists in those sectors is likely to be higher than it would otherwise have been. In addition, the potential for off-shoring would drive up the derived elasticity of demand for computer scientists. However, even if, for these reasons, the derived demand for computer sciences in computer manufacturing and computer services was quite high, a small enough share of computer scientists work in these industries, that it is hard to believe either agglomeration effects or off-shoring can drive up the derived demand elasticity for computer scientists that much. Additionally, if it would have been easy for employers to outsource, CEOs like Microsoft's Bill Gates would not have been lobbying to increase the H-1B visa cap. It is hard to reconcile the fact that the computer industry is lobbying so hard for easier access to foreigners, if it did not matter where their workforce was located.

[37] According to the Census, roughly 30% of computer scientists worked in either the computer manufacturing or the computer services three-digit industries during 1990 and 2000.

Traditionally, exogenous shifts in supply are used to identify demand curves. In our case, while there is a plausibly exogenous component to the increased representation of the foreign born amongst the US Science and Engineering workforce, our visa system ensures that there is a large endogenous component. In theory, it might be possible to get some leverage on identifying the labor demand curve for CS workers by comparing the results of the counterfactual simulation for the different $\gamma's$ to the observations of what happened in the Information Technology (IT) sector in the the mid 1970s. Specifically, as described in Bound et al. (2013), during this earlier period, the IT sector experienced a significant transformation due to the introduction of the microprocessor, which generated an increase in the demand for IT workers. However, firms had substantially less access to foreign labor during that earlier boom than they did during the 1990s. This happened because there was a sharp increase in the supply of college graduates from overseas in the past decades, but also due to the change in the US visa system in the early 1990s that facilitated a greater inflow of high skilled foreigners via employer-sponsored visas.

Our strategy would be to use our calibrated model to simulate what would happen if firms had less access to foreign high-skilled labor in the 1990s boom and compare these simulations to the earlier boom. Comparisons between simulation results with different values of γ and what actually happened earlier would help narrow plausible values for γ. Intuitively, if demand is relatively elastic, the loss of access to foreigners would have relatively little impact on wages, but a large impact on total CS employment. Whereas a less elastic demand curve would have a large effect on wages and less of an effect on total CS employment. This kind of exercise is valid only under the strong assumptions that our economic model accurately reflects that labor market for IT workers, and that the demand and supply elasticities were the same during the two periods and that the two shocks generated shifts in the labor demand of roughly the same magnitude. However heroic such assumptions might be, the strategy fails for a simpler reason. The strategy requires comparing wage and employment changes for a small segment of the workforce across periods. Our estimates were simply not reliable enough for such exercises to be meaningful.

Given the data limitations and other complications discussed in this section, we cannot provide an estimate for the value of γ, but our discussion suggests that the elasticity of demand for computer scientists should not be too large and that the values presented in this paper cover a plausible range.

3.5 Discussion

The model we have developed in this paper suggests an intermediate position as the most reasonable one in the debate over the effects of high-skilled immigration, on US workers. Focusing on the tech boom of the 1990s, we develop a model that allows us to answer the counterfactual question: what would have happened to overall employment, to the employment of US residents, and to wages in the IT sector had the immigration of computer scientists been restricted to its level as of the early 1990s before the tech boom? Our results suggest a middle ground between the two sides of this debate.

First, our estimates suggest that even without foreign trained computer scientists, the supply of computer scientists to the US market is quite elastic, especially in the medium run, as the students induced to study computer science by the increased opportunities in the field begin to enter the market. For elasticities of demand that lie between -1.3 and -4.0, we show that had firms not been able to hire immigrants through the late 1990s, the wages of US trained computer scientists would have been 2.8% to 3.8% higher than they were, and there would have been 7% to 13.6% more Americans working as computer scientists.

At the same time our estimates suggest that were it not for the immigrant computer scientists that firms were able to hire, the growth in the number of computer scientists in the economy would have been significantly slowed. Our estimates suggest that total employment in the CS sector would have been 3.8-9% lower if firms were not able to hire additional immigrants during the late 1990s, thus implying that the fact that firms could hire immigrants during the 1990s increased output and lowered both input and output prices in the computer scientist intensive sectors of the economy. How much these developments benefited stock holders and consumers depends on the nature of the output market, which we have not tried to model. The increased employment of computer scientists would also have increased the demand for complementary production inputs, such as software marketing and sales workers. Furthermore, the availability of foreign CS workers made the CS labor supply curve more elastic, further enhancing this demand increase for complements.

Under the assumption that the tech boom of the 1990s exogenously increased the demand for computer scientists, we have been able to reliably estimate supply curves. Estimating the slope of the labor demand curve was substantially more difficult. In other contexts, labor economists have been willing to assume some degree of exogeneity to immigrant supplies. In the current framework, the institutional context implies that immigrant CS labor is completely endogenous to labor demand.

While we cannot reliably estimate the slope of the demand curve for computer scientists, we believe that we can reject any notion that the demand curve for computer scientists is close to perfectly elastic. Perfectly elastic demand curves are inconsistent with the rising wages for computer scientists that we observe during the 1990s. As long as the demand curve for computer scientists is downward sloping, the increased access employers had to foreign-trained, skilled immigrants during the 1990s works to lower both the wages and employment opportunities for US trained computer scientists.

Our paper should be viewed as a first-step towards modeling the US labor market for computer-scientists. In the model we incorporate features that were ignored in earlier models developed by Freeman (1976) and Ryoo and Rosen (2004a). Specifically we model both the possibility that individuals might switch occupations, and the possibility that firms might hire immigrants from abroad. In the context of computer scientists both are clearly important. We focused on the market for computer scientists. In the context of other scientific fields where a masters or PhD are essential, it would also be important to model foreign participation in US graduate programs as well. Such an effort would need to model both the demand for and supply of higher education. While we believe that such an effort would be of considerable value, we leave it for future research.

Table 3.1: Fraction of Computer Scientists and Immigrants in the US Workforce by Occupation

Year	1970	1980	1990	2000	2010
Computer Scientists:					
as a fraction of workers with a BA/MA	1.68%	1.83%	3.30%	5.66%	5.28%
as a fraction of STEM college graduates	16.86%	23.60%	35.99%	53.31%	54.90%
Immigrants:					
as a fraction of Bachelor's/Master's	2.10%	5.43%	6.86%	8.41%	12.77%
as a fraction of Computer Scientists	2.37%	7.09%	11.06%	18.59%	27.82%
as a fraction of Other STEM workers	3.63%	9.72%	10.71%	12.69%	18.21%

Note: Sample restricted to employed workers with a Bachelors or a Masters degree. Definition of Computer Scientists and STEM workers determined by occupational coding (for details see Data Appendix online). Immigrant is defined as one born abroad, and migrated to the US after the age of 18.

Source: US Census (years 1970 to 2000); ACS (2010)

Table 3.2: Calibrated Parameters

Parameter	Description	γ		
		0.25	0.50	0.75
Calibrated Parameters		Calibrated Value		
α_0	Mean taste for not studying CS	0.0940	0.0943	0.0836
σ_0	Std. dev. of study area taste shocks	0.0001	0.0001	0.0002
α_1	Mean taste for not working in CS	0.3715	0.3486	0.3673
σ_1	Std. dev. of occupation taste shocks	0.1385	0.1364	0.1439
c_{R1}	Foreign linear recruitment cost	0.5247	0.5228	0.5221
c_{R2}	Foreign quadratic recruitment cost	0.0102	0.0109	0.0124
λ_0	Sector switching constant cost	0.1159	0.1164	0.1031
λ_1	Sector switching linear cost	0.0138	0.0119	0.0151
λ_2	Sector switching quadratic cost	0.0006	0.0004	0.0003
Economic Results				
$\epsilon_{L_d,w}$	Long run effective labor demand elasticity	1.33	2.00	4.00
$\epsilon_{L_s,w}$	Long run US occupational labor supply elasticity	5.4612	5.5743	5.3404
$\epsilon_{p,w}$	Long run US CS enrollment elasticity	11.6954	11.2624	11.7071
$\epsilon^s_{L_s,w}$	Short run US occupational labor supply elasticity	0.5591	0.6745	0.6642
$\epsilon^s_{p,w}$	Short run US CS enrollment elasticity	10.2834	11.3758	11.3386
AC_F	Average cost of recruiting foreign worker	0.5312	0.5299	0.53

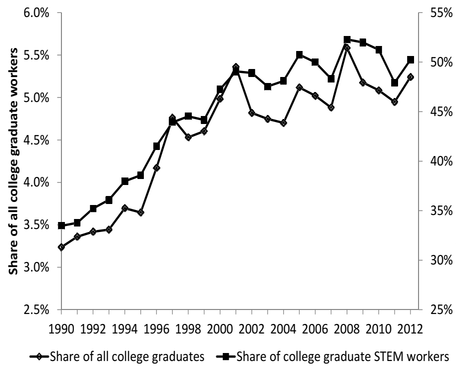
Note: The average cost of recruiting a foreign worker is measured in units of average annual US non-CS STEM worker wages. The parameter γ determines the labor demand elasticity to wages.

Table 3.3: Summary of Results from Counterfactual Simulation

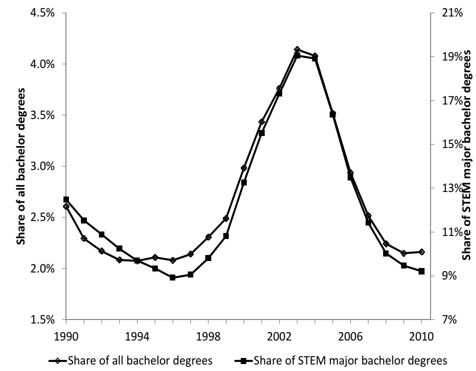
% Differences between Simulated Economy Holding F Constant and Actual Outcomes in 2004			
Variable	γ		
	0.25	0.5	0.75
CS Wages	3.8%	3.2%	2.8%
CS US Native Employment	13.6%	12.5%	7.0%
CS Enrollment	25.5%	20.2%	19.9%
Total Employment	-3.8%	-4.6%	-9.0%

Note: The counterfactual simulates an economy from 1994-2009 in which the level of foreign CS workers is not allowed to increase from its 1994 value. The parameter γ determines the labor demand elasticity to wages. See section 4 for details.

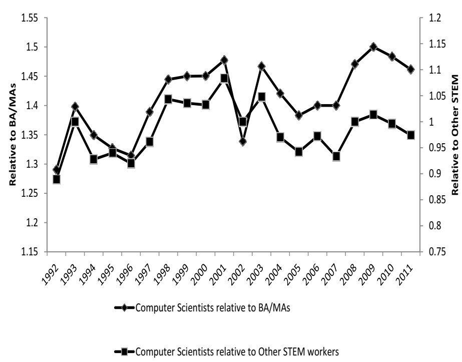
Figure 3.1: Major Trends (1990 to 2012)



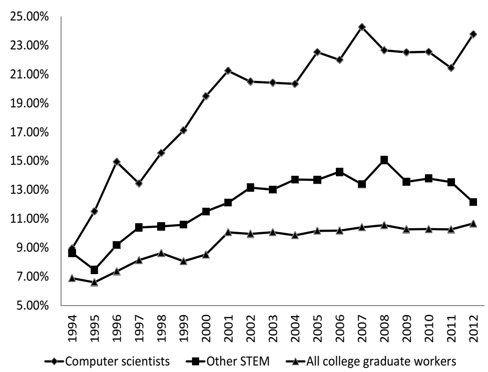
(a) Fraction of Computer Scientists in US Workforce



(b) Computer Science Fraction of Bachelor Degrees Awarded in US



(c) Relative Earnings of Computer Scientists

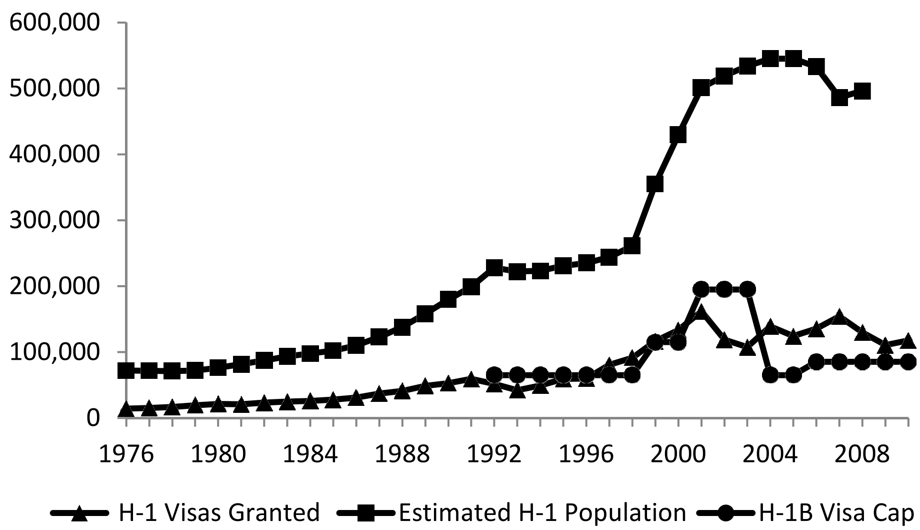


(d) Foreign Born and Immigrated at Age 18 or Older Fraction of Employed Population by Occupation

Note: Sample restricted to employed workers with a Bachelors or a Masters degree. Definition of Computer Scientists and STEM workers determined by occupational coding (for details see Data Appendix online). STEM majors are defined as engineering, computer and math sciences and natural science. Earning are median weekly earnings. Imputed values excluded, and values are lagged by one year due to retrospective nature of the survey. Immigrant defined as one born abroad, and migrated to the US after the age of 18. Immigration status is not available in the CPS before 1994.

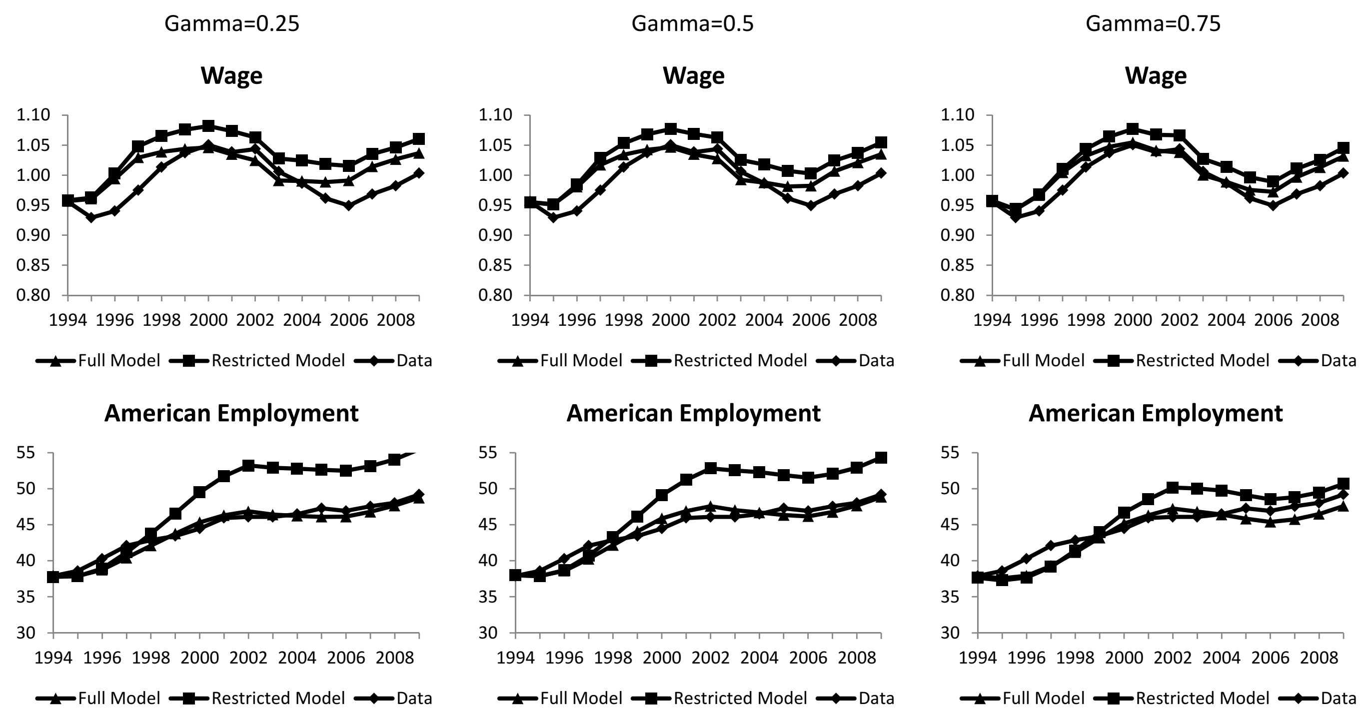
Sources: March CPS (for employment, earnings, and immigrants); IPEDS Completions Survey (for degrees)

Figure 3.2: H-1 and H-1B Visa Population



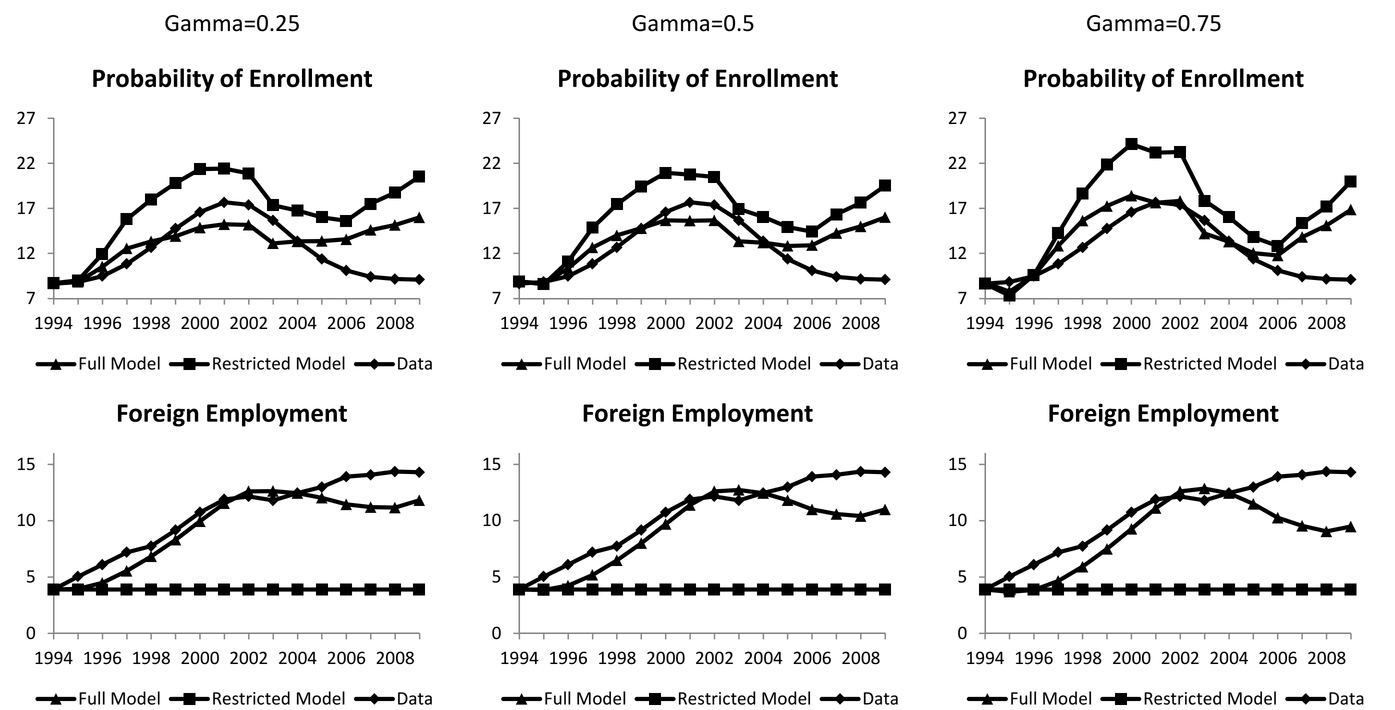
Note: Population stock is constructed using estimations of inflow (visas granted) and outflow (deaths, permanent residency, or emigration) of H-1 workers. In later years, the number of visas granted could exceed the visa cap due to exemptions for foreigners who work at universities and non-profit research facilities.

Figure 3.3: Model and Counterfactual (1/2)



Note: The full model is the simulation of the economy using the calibrated parameters. The restricted model simulates the same calibrated model, but restricting firms to keep their foreign temporary worker CS employment to its 1994 value. Wages are relative to other STEM occupations. Employment and enrollment are shares of STEM workers and undergrad STEM enrollment, respectively, and are multiplied by 100. The parameter gamma determines the labor demand elasticity to wages. See Section 4 for details.

Figure 3.4: Model and Counterfactual (2/2)



Note: The full model is the simulation of the economy using the calibrated parameters. The restricted model simulates the same calibrated model, but restricting firms to keep their foreign temporary worker CS employment to its 1994 value. Wages are relative to other STEM occupations. Employment and enrollment are shares of STEM workers and undergrad STEM enrollment, respectively, and are multiplied by 100. The parameter gamma determines the labor demand elasticity to wages. See Section 4 for details.

BIBLIOGRAPHY

Abrams, David S. and Albert H. Yoon, "The Luck of the Draw: Using Random Case Assignment to Investigate Attorney Ability," *University of Chicago Law Review*, 2007, *74* (4), 1145–1177.

Agrawal, Ajay, John Horton, Nico Lacetera, and Elizabeth Lyons, "Digitization and the Contract Labor Market: A Research Agenda," in "Economics of Digitization," University of Chicago Press, 2013.

Altonji, Joseph, Todd Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 2005, *113* (1), 151–184.

Andrews, Donald W. K. and Marcia M. a. Schafgans, "Semiparametric Estimation of the Intercept of a Sample Selection Model," *Review of Economic Studies*, July 1998, *65* (3), 497–517.

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

Bechtold, Stefan and Catherine Tucker, "Trademarks, Triggers and Online Search," *Available at SSRN 2266945*, 2014.

Blake, Thomas, Chris Nosko, and Steven Tadelis, "Consumer heterogeneity and paid search effectiveness: A large scale field experiment," *NBER Working Paper*, 2014, p. 126.

Blundell, Richard W and James L. Powell, "Endogeneity in Semiparametric Binary Response Models," *The Review of Economic Studies*, 2004, *71* (3), 655–679.

Boal, William M. and Michael R Ransom, "Monopsony in the Labor Market," *Journal of Economic Literature*, 1997, *35* (1), pp. 86–112.

Bollinger, Christopher and Barry Hirsch, "How Well are Earnings Measured in the Current Population Survey? Bias from Nonresponse and Proxy respondents," in "North American Summer Meetings of the Econometric Society, Duke University" 2007.

Bonhomme, Stephane and Gregory Jolivet, "The pervasive absence of compensating differentials," *Journal of Applied Econometrics*, 2009, *24* (5), 763–795.

Borjas, George, "The Economic Analysis of Immigration," *Handbook of Labor Economics*, 1999, *Vol 3A, edited by Orley Ashenfelter and David Card, North-Holland.*

_ , "Immigration in High-Skill Labor Markets: The Impact of Foreign Students on the Earnings of Doctorates," *Science and Engineering Careers in the United States: An Analysis of Markets and Employment*, 2009, *Richard B. Freeman and Daniel L. Goroff, eds.*, 131 – 161.

Bound, John, Breno Braga, Joseph Golden, and Sarah Turner, "Pathways to Adjustment: The Case of Information Technology Workers," *American Economic Review: Papers and Proceedings*, 2013.

Brown, Charles, "Equalizing Differences in the Labor Market," *The Quarterly Journal of Economics*, 1980, *94* (1), 113–134.

Campbell-Kelly, M. and W. Aspray, *Computer: A History of the Information Machine* The Sloan Technology Series, Westview Press, 2004.

Carrington, William J., "The Alaskan Labor Market during the Pipeline Era," *Journal of Political Economy*, 1996, *104* (1), pp. 186–218.

Chamberlain, Gary, "Asymptotic efficiency in semi-parametric models with censoring," *Journal of Econometrics*, 1986, *32*, 189–218.

Clemens, Michael A, "Why Do Programmers Earn More in Houston Than Hyderabad? Evidence from Randomized Processing of US Visas," *American Economic Review, Papers & Proceedings*, 2013, *103(3): 198-202.*

Collard, Fabrice and Michel Juillard, "A Higher-Order Taylor Expansion Approach to Simulation of Stochastic Forward-Looking Models with an Application to a Non-Linear Phillips Curve," *Computational Economics*, 2001a, *17(2)*, 125–139.

_ **and** _ , "Accuracy of Stochastic Perturbation Methods: The Case of Asset Pricing Models.," *Journal of Economic Dynamics and Control*, 2001b, *25(7)*, 979–999.

Cosslett, Stephen R, "Semiparametric estimation of a regression model with sample selectivity," in "Nonparametric and semiparametric methods in econometrics and statistics," Cambridge University Press: Cambridge, UK, 1991, pp. 175—-97.

Dehejia, Rajeev H. and Sadek Wahba, "Propensity score-matching methods for nonexperimental causal studies," *Review of Economics and statistics*, 2002, *84* (1), 151–161.

Delfgaauw, Josse and Robert Dur, "Managerial talent, motivation, and self-selection into public management," *Journal of Public Economics*, October 2010, *94* (9-10), 654–660.

Denning, Peter J., "The Computing Field: Structure," in "Wiley Encyclopedia of Computer Science and Engineering" 2008.

DeVaro, J. and Dana Brookshire, "Promotions and incentives in nonprofit and for-profit organizations," *Industrial and Labor Relations Review*, 2006, *60* (3), 311–339.

Duan, Naihua, "Smearing estimate: a nonparametric retransformation method," *Journal of the American Statistical Association*, 1983, *78* (383), 605–610.

Dynarski, Susan M. and Judith E. Scott-Clayton, "The Cost of Complexity in Federal Student Aid: Lessons from Optimal Tax Theory and Behavioral Economics," *National Tax Journal*, 2006, *LVIV* (2).

Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz, "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords," *American Economic Review*, March 2007, *97* (1), 242–259.

Field, Erica, "Educational Debt Burden and Career Choice: Evidence from a Financial Aid Experiment at NYU Law School," *American Economic Journal: Applied Economics*, January 2009, *13* (2), 1–21.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 2008, *98* (1), 394–425.

Frank, Robert H., "What price the moral high ground?," *Southern Economic Journal*, July 1996, *63* (1), 1–17.

Freeman, Richard B, "Supply and Salary Adjustments to the Changing Science Manpower Market: Physics, 1948-1973," *American Economic Review*, March 1975, *65* (1), 27–39.

Freeman, Richard B., "A Cobweb Model of the Supply and Starting Salary of New Engineers," *Industrial and Labor Relations Review*, 1976, *29* (2), pp. 236–248.

_ , "An Empirical Analysis of the Fixed Coefficient "Manpower Requirements" Model, 1960-1970," *The Journal of Human Resources*, 1980, *15* (2), pp. 176–199.

Freeman, Richard B, "Does Globalization of the Scientific/Engineering Workforce Threaten U.S. Economic Leadership?," *Innovation Policy and the Economy*, 2006, *Volume 6, Adam B. Jaffe, Josh Lerner, and Scott Stern.*

Freeman, Richard B., "What Does Global Expansion of Higher Education Mean for the US?," Working Paper 14962, National Bureau of Economic Research May 2009.

French, Eric and Christopher Taber, "Identification of Models of the Labor Market," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Amsterdam: Elsevier, November 2011, pp. 537–617.

Gervais, Daniel, Martin Holmes, Paul Kruse, Glenn Perdue, and Caprice Roberts, "Is Profiting from the Online Use of Another's Property Unjust? The Use of Brand Names as Paid Search Keywords," *IDEA: The IP Law Review*, 2013, *53*, 131.

Goddeeris, John H., "Compensating Differentials and Self-Selection: An Application to Lawyers," *Journal of Political Economy*, January 1988, *96* (2), 411–428.

Hamermesh, Daniel S., "Labor Demand and the Structure of Adjustment Costs," *The American Economic Review*, 1989, *79* (4), pp. 674–689.

Handy, Femida and Eliakim Katz, "The Wage Differential between Nonprofit Institutions and Corporations: Getting More by Paying Less?," *Journal of Comparative Economics*, June 1998, *26* (2), 246–261.

Hansmann, Henry B., "The Role of Nonprofit Enterprise," *Yale Law Journal*, 1980, *89* (5), 835–901.

Heckman, James, "Varieties of selection bias," *American Economic Review*, 1990, *80* (2).

Heckman, James J., "Sample selection bias as a specification error," *Econometrica*, 1979, *1* (11), 153–161.

Hershbein, Brad, "Worker Signals Among New College Graduates: The Role of Selectivity and GPA," Working Paper 2011.

Heyes, Anthony, "The economics of vocation or 'why is a badly paid nurse a good nurse'?," *Journal of Health Economics*, May 2005, *24* (3), 561–9.

Horton, John J., "Online labor markets," in "Internet and Network Economics," Vol. 6484 of *Lecture Notes in Computer Science*, Springer, 2010, pp. 515–522.

Hunt, Jennifer, "Which Immigrants Are Most Innovative and Entrepreneurial? Distinctions by Entry Visa," *Journal of Labor Economics*, 2011, *29* (3), 417 – 457.

_ , "Does the United States Admit the Best and Brightest Computer and Engineering Workers?," *Working Paper*, 2012.

_ **and Marjolaine Gauthier-Loiselle**, "How Much Does Immigration Boost Innovation?," *American Economic Journal: Macroeconomics*, 2010, *2(2)*, 31–56.

Hurst, Erik and Benjamin W. Pugsley, "What Do Small Businesses Do?," *Brookings Papers on Economic Activity*, 2011, (Fall), 73–118.

Imbens, Guido and Jeffrey Wooldridge, "What's New In Econometrics? Lecture 6," Lecture Notes, National Bureau of Economic Research 2007.

Jansen, Bernard J., Zhe Liu, and Zach Simon, "The effect of ad rank on the performance of keyword advertising campaigns: The Effect of Ad Rank on Performance of Keyword Advertising Campaigns," *Journal of the American Society for Information Science and Technology*, October 2013, *64* (10), 2115–2132.

Juillard, Michel, “DYNARE: A Program for the Resolution and Simulation of Dynamic Models with Forward Variables through the Use of a Relaxation Algorithm.,” *CEPREMAP*, 1996, *9602*.

Kambourov, Gueorgui and Iourii Manovskii, “Occupational Mobility and Wage Inequality,” *The Review of Economic Studies*, 2009, *76* (2), 731–759.

__ **and** __ , “Occupational Specificity Of Human Capital,” *International Economic Review*, 02 2009, *50* (1), 63–115.

Katz, Lawrence F and Kevin M Murphy, “Changes in Relative Wages, 1963-1987: Supply and Demand Factors,” *The Quarterly Journal of Economics*, February 1992, *107* (1), 35–78.

Kerr, Sari Pekkala, William R. Kerr, and William F. Lincoln, “Skilled Immigration and the Employment Structures of U.S. Firms,” *Working Paper*, 2013.

Kerr, William and William Lincoln, “The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Invention,” *Journal of Labor Economics*, 2010, *28(3)*, 473–508.

Kirkegaard, Jacob, “Outsourcing and Skill Imports: Foreign High-Skilled Workers on H-1B and L-1 Visas in the United States,” *Working Paper no. 05-15, Peterson Institute for International Economics,Washington, DC.*, 2005.

Kline, Patrick, “Understanding Sectoral Labor Market Dynamics: An Equilibrium Analysis of the Oil and Gas Field Services Industry,” Cowles Foundation Discussion Papers 1645, Cowles Foundation for Research in Economics, Yale University 2008.

Krugman, Paul, “A Model of Innovation, Technology Transfer, and the World Distribution of Income,” *Journal of Political Economy*, 1979, *Vol. 87, No. 2*, 253–266.

Kydland, Finn and Edward Prescott, “Time to Build and Aggregate Fluctuations,” *Econometrica*, 1982, *50(6)*, 1345–1370.

Landers, Renee M., James B. Rebitzer, and Lowell J. Taylor, “Rat race redux: Adverse selection in the determination of work hours in law firms,” *The American Economic Review*, 1996, *86* (3), 329–348.

Lang, Kevin and Sumon Majumdar, “the Pricing of Job Characteristics When Markets Do Not Clear: Theory and Policy Implications,” *International Economic Review*, November 2004, *45* (4), 1111–1128.

Leete, Laura, “Wage equity and employee motivation in nonprofit and for-profit organizations,” *Journal of Economic Behavior and Organization*, December 2000, *43* (4), 423–446.

__ , “Whither the Nonprofit Wage Differential? Estimates from the 1990 Census,” *Journal of Labor Economics*, January 2001, *19* (1), 136–170.

Lehmann, Jee-Yeon K., “Job Assignment and Promotion Under Statistical Discrimination : Evidence from the Early Careers of Lawyers,” Working Paper, University of Houston 2010.

Leiner, Barry M., Vinton G. Cerf, David D. Clark, Robert E. Kahn, Leonard Kleinrock, Daniel C. Lynch, Jon Postel, Lawrence G. Roberts, and Stephen S. Wolff, "The past and future history of the Internet," *Commun. ACM*, February 1997, *40* (2), 102–108.

Lemieux, Thomas, "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *The American Economic Review*, 2006, *96*, 461–498.

Lewis, Randall A. and Justin M. Rao, "On the near impossibility of measuring the returns to advertising," *Unpublished manuscript*, 2013.

_ , _ , and David H. Reiley, "Here, there, and everywhere: correlated online behaviors can lead to overestimates of the effects of advertising," in "Proceedings of the 20th international conference on World wide web" ACM 2011, p. 157166.

Lowell, B. Lindsay, "H-1B Temporary Workers: Estimating the Population," Working Paper 12, The Center for Comparative Immigration Studies 2000.

Lowell, B Lindsay, "Skilled temporary and permanent immigrants in the United States," *Population Research and Policy Review*, 2001, *20* (1-2), 33–58.

Lucas, Robert E.B., "Hedonic wage equations and psychic wages in the returns to schooling," *The American Economic Review*, 1977, *67* (4), 549–558.

Manning, Alan, *Monopsony in motion: imperfect competition in labor markets*, Princeton University Press, Oxford, UK. ISBN 0691113122, 2003.

Matloff, Norman, "On the Need for Reform of the H-1B Non-Immigrant Work Visa in Computer-Related Occupations," *University of Michigan Journal of Law Reform*, 2003, *36, no. 4*, 815–914.

_ , "Needed Reform for the H-1B and L-1 Work Visas (and Relation to Offshoring)," *Working Paper, University of California, Davis.*, 2004.

Miano, John, "The Bottom of the Pay Scale: Wages for H-1B Computer Programmers.," *Center for Immigration Studies, Washington, DC.*, 2005.

_ , "H-1B Visa Numbers: No Relationship to Economic Need," *Center for Immigration Studies, Washington, DC.*, 2008.

Mithas, Sunil and Henry C. Lucas, "Are Foreign IT Workers Cheaper? U.S. Visa Policies and Compensation of Information Technology Professionals," *Management Science*, May 2010, *56* (5), 745–765.

Mocan, H. Naci and Erdal Tekin, "Nonprofit sector and part-time work: an analysis of employer-employee matched data of child care workers," *The Review of Economics and Statistics*, 2003, *85* (February), 38–50.

Mukhopadhyay, Sankar and David Oxborrow, "The Value of an Employment-Based Green Card," *Demography*, 2012, *49*, 219–237.

Nathan, Max, "The Wider Economic Impacts of High-Skilled Migrants: A Survey of the Literature," *IZA Working Papers*, 2013, *Discussion Paper No. 7653.*

Newey, Whitney K., "Two-step series estimation of sample selection models," *Econometrics Journal*, January 2009, *12*, S217–S229.

Oreopoulos, Philip and Kjell Salvanes, "Priceless: The Nonpecuniary Benefits of Schooling," *Journal of Economic Perspectives*, 2011, *25* (1), 159–84.

Oyer, Paul and Scott Schaeffer, "The Returns to Attending a Prestigious Law School," Working Paper, Stanford University 2010.

Peri, Giovanni and Chad Sparber, "Highly-Educated Immigrants and Native Occupational Choice," *Industrial Relations*, 2011, *Vol. 50 (3)*, 385–411.

_ , Kevin Shih, and Chad Sparber, "STEM Workers, H1B Visas and Productivity in US Cities," 2013.

Preston, Anne E., "The Nonprofit Worker in a For-Profit World," *Journal of Labor Economics*, January 1989, *7* (4), 438–463.

Rebitzer, James B. and Lowell Taylor, "Efficiency wages and employment rents: The employer-size wage effect in the job market for lawyers," *Journal of Labor Economics*, January 1995, *13* (4), 678.

Roback, Jennifer, "Wages, rents and the quality of life," *Journal of Political Economy*, 1982, *90* (6), 1257–1278.

Romer, Paul M, "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, 1986, *Vol. 94, No. 5*, 1002–1037.

Rosen, Sherwin, "The Theory of Equalizing Differences," in Orley Ashenfelter and Richard Layard, eds., *Handbook of Labor Economics*, Amsterdam: Elsevier, November 1986, pp. 641–692.

Rothstein, Jesse and Cecilia E. Rouse, "Constrained after college: Student loans and early-career occupational choices," *Journal of Public Economics*, February 2011, *95* (1-2), 149–163.

Roy, A.D., "Some thoughts on the distribution of earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

Ruhm, Christopher J. and Carey Borkoski, "Compensation in the Nonprofit Sector," *The Journal of Human Resources*, January 2003, *38* (4), 992–1021.

Rust, John, "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 1987, *55* (5), pp. 999–1033.

Ryoo, Jaewoo and Sherwin Rosen, "The Engineering Labor Market," *Journal of Political Economy*, February 2004, *112* (S1), S110–S140.

— **and** — , "The Engineering Labor Market," *Journal of Political Economy*, February 2004, *112* (S1), S110–S140.

Sargent, Thomas J., "Estimation of Dynamic Labor Demand Schedules under Rational Expectations," *Journal of Political Economy*, 1978, *86* (6), pp. 1009–1044.

Sauer, Robert M., "Job mobility and the market for lawyers," *Journal of Political Economy*, February 1998, *106* (1), 147–171.

Sayedi, Amin, Kinshuk Jerath, and Kannan Srinivasan, "Competitive Poaching in Sponsored Search Advertising and Its Strategic Impact on Traditional Advertising," *Marketing Science*, July 2014, *33* (4), 586–608.

Schafgans, Marcia M.a., "Gender wage differences in Malaysia: parametric and semiparametric estimation," *Journal of Development Economics*, December 2000, *63* (2), 351–378.

Shapiro, Matthew D., "The Dynamic Demand for Capital and Labor," *The Quarterly Journal of Economics*, 1986, *101* (3), pp. 513–542.

shin Hwang, Hae, Dale T. Mortensen, and W. Robert Reed, "Hedonic Wages and Labor Market Search," *Journal of Labor Economics*, 1998, *16* (4), 815–847.

— **, W. Robert Reed, and Carlton Hubbard**, "Compensating Wage Differentials and Unobserved Productivity," *Journal of Political Economy*, 1992, *100* (4), 835–858.

Silva, Joao Santos and Silvana Tenreyro, "The log of gravity," *The Review of Economics and Statistics*, 2006, *88* (4), 641–658.

Smith, Jeffrey A. and Petra E. Todd, "Does matching overcome LaLonde's critique of nonexperimental estimators?," *Journal of Econometrics*, March 2005, *125* (1-2), 305–353.

Stern, Scott, "Do Scientists Pay to Be Scientists?," *Management Science*, June 2004, *50* (6), 835–853.

Stewart, Jay, "Recent Trends in Job Stability and Job Security: Evidence from the March CPS," *BLS Working Papers, Office of Employment and Unemployment Statistics, Working paper 356*, 2002.

Stinebrickner, Todd R. and Ralph Stinebrickner, "Math or Science? Using Longitudinal Expectations Data to Examine the Process of Choosing a College Major," Working Paper 16869, National Bureau of Economic Research 2011.

Stock, James, Jonathan Wright, and Motohiro Yogo, "A survey of weak instruments and weak identification in generalized method of moments," *Journal of Business and Economic Statistics*, October 2002, *20* (4), 518–529.

Sullivan, Paul and Ted To, "Job Dispersion and Compensating Wage Differentials," Lecture Notes, National Bureau of Economic Research 2012.
Temple University Beasley School of Law: Where Our Grads Have Gone

***Temple University Beasley School of Law: Where Our Grads Have Gone*.**

U.S. Department of Commerce, *"Digital Economy 2000," Technical Report, U.S. Department of Commerce 2000.*

U.S. General Accounting Office, *"Immigration and the Labor Market: Nonimmigrant Alien Workers in the United States," Technical Report GAO/PEMD-92-17., U.S. General Accounting Office, Washington, D.C. 1992.*

_ , *"H-1B Foreign Workers: Better Controls Needed to Help Employers and Protect Workers," Technical Report GAO/HEHS 00-157., U.S. General Accounting Office, Washington, D.C. 2000.*

_ , *"H-1B Visa Program: Reforms Are Needed to Minimize the Risks and Costs of Current Program," Technical Report GAO/HEHS 00-157., U.S. General Accounting Office, Washington, D.C. 2011.*

U.S. Immigration and Naturalization Service, *"Characteristics of Specialty Occupation Workers (H 1B)," Technical Report, U.S. Immigration and Naturalization Service, Washington, D.C February 2000.*

Varian, Hal R., *"Position auctions," International Journal of Industrial Organization, December 2007, 25 (6), 1163–1178.*

_ **and Christopher Harris**, *"The VCG Auction in Theory and Practice ," American Economic Review, May 2014, 104 (5), 442–445.*

Vella, Francis, *"Estimating models with sample selection bias: a survey," Journal of Human Resources, 1998, 33 (1), 127–169.*

Weisbrod, Burton A., *"Nonprofit and Propriety Sector Behavior: Wage Differentials Among Lawyers," Journal of Labor Economics, 1983, 1 (3), 246–263.*

Wooldridge, Jeffrey, *Econometric Analysis of Cross Section and Panel Data, MIT Press, 2002.*