

THREE ESSAYS IN LABOR AND EDUCATION ECONOMICS

by

Breno Braga

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, Co-Chair

Professor Charles Brown, Co-Chair

Professor David Lam

Professor Brian McCall

© Breno Braga 2014

DEDICATION

Aos meus pais.

ACKNOWLEDGMENT

I am very grateful to my dissertation committee, John Bound, Charlie Brown, David Lam and Brian McCall, for all the comments and suggestions that improved this dissertation significantly. I am especially thankful to John and Charlie for the many hours of discussion and endless constructive comments on my work. Both of them have hugely influenced the way I think about economics and conduct my research.

The second chapter of this dissertation is a joint work with Paola Bordon and the third chapter is a joint work with John Bound, Joe Golden and Gaurav Khanna. I am very grateful to all these coauthors from whom I have learned much throughout our time together.

I wish to thank all my colleagues in the Economics PhD program, in particular Aditya Aladangady, Vanessa Alviarez, Lasse Brune, Sebastian Calonico, Italo Gutierrez, Monica Hernandez, Cynthia Doniger and Peter Hudomiet. They were an incredible source of academic and personal support and I deeply appreciate all their suggestions.

I am continually thankful for my beautiful girlfriend, Katherine Axelsen. In addition to proofreading all my papers, her kindness and love provided great happiness to me these past 4 years. I could not have a better partner to take the next step of my life.

I am very grateful to my family in Brazil, especially my sister Tássia, grandmothers Therezinha and Dadá, and all my uncles and aunts. Despite being so far, their moral support and

love were always present during my work and crucial to all my accomplishments.

Last, but definitely not least, I would like to deeply thank my parents Nize and Antonio Braga. Since an early age they have been obsessed with providing my sister and I with the best possible education despite many adverse situations. This dissertation is nothing but the successful result for their tireless efforts.

TABLE OF CONTENTS

DEDICATION	ii
ACKNOWLEDGMENT	iii
LIST OF TABLES	viii
LIST OF FIGURES	x
Chapter I. Schooling, Experience, Career Interruptions, and Earnings	1
1.1 Introduction	1
1.2 Potential Experience and Career Interruptions in the Mincer Equation	6
1.3 Empirical Dynamics	11
1.3.1 Data	11
1.3.2 Earnings Dynamics Estimation	14
1.4 Model	27
1.4.1 The model environment	28

1.4.2	Equilibrium	33
1.4.3	Some Extensions	40
1.5	Conclusion	42
1.6	Theory Appendix	43
1.6.1	Derivation of $\mu^*(x, u)$ and $\tilde{\mu}$	43
1.6.2	Proof of Proposition 3 and 4	44

Chapter II. Employer Learning, Statistical Discrimination and University

Prestige		65
2.1	Introduction	65
2.2	Institutional Framework	70
2.3	Data	73
2.4	Regression Discontinuity Test	75
2.4.1	Employer Learning Statistical Discrimination Model	75
2.4.2	The Admission Process and the RD Design	82
2.4.3	Results	85
2.5	Conclusion	88

Chapter III. Recruitment of Foreigners in the Market for Computer Scientists

in the US		97
3.1	Introduction	97

3.2	The Market for Computer Scientists in the 1990s	102
3.2.1	The Information Technology Boom of the Late 1990s	102
3.2.2	The Immigrant Contribution to the Growth of the High Tech Workforce	103
3.2.3	The Previous Literature on the Impact of Immigrants on the High Tech Workforce in the US	107
3.3	A Dynamic Model of Supply and Demand of Computer Scientists	110
3.3.1	Labor Supply of American Computer Scientists	111
3.3.2	Labor Supply of Foreign Computer Scientists	117
3.3.3	Labor Demand for Computer Scientists	118
3.3.4	Equilibrium	122
3.4	Calibration and Simulation	124
3.4.1	Identification and Calibration Method	124
3.4.2	Calibration results	127
3.4.3	Simulation of Fixed Foreign Worker Population Counterfactual	129
3.4.4	Identification of Labor Demand	131
3.5	Discussion	133

Bibliography	143
---------------------	------------

LIST OF TABLES

1.1	Literature Review	46
1.2	Work History Variables	47
1.3	Descriptive Statistics	48
1.4	The Effect of Schooling, Experience, and Career Interruptions on Earnings .	49
1.5	The Effect of Schooling, Experience, and Career Interruptions on Earnings, Other Demographic Groups	50
1.6	The Effect of Schooling, Experience, and Career Interruptions on Earnings - Leaving School Year as First Year a Responded Left School	51
1.7	The Effect of Unemployment and Schooling on Earnings by Timing of Unem- ployment	52
1.8	The Effect of Schooling, Experience, and Career Interruptions on Earnings, Individual Fixed Effect	53
2.9	Descriptive Statistics for Selective and Non-Selective Universities	89
2.10	Earnings for Selective and Non-Selective Universities	89
2.11	EL-SD Regression Discontinuity Test	90

2.12 EL-SD Regression Discontinuity Test - Robustness Checks	91
3.13 Fraction of Computer Scientists and Immigrants in the US Workforce by Occupation	136
3.14 Calibrated Parameters	137
3.15 Summary of Results from Counterfactual Simulation	138

LIST OF FIGURES

1.1	Employment Attachment over the Life-Cycle - High School Graduates	54
1.2	Employment Attachment over the Life-Cycle - BA or More Graduates	55
1.3	Earnings Coefficient on Schooling by Work Experience Groups	56
1.4	Earnings Coefficient on Schooling by Cumulative Years Unemployment Groups	57
1.5	Earnings Coefficient on Schooling by Cumulative Years OLF Groups	58
1.6	Log Earnings - Work Experience Profile	59
1.7	Log Earnings - Cumulative Years Unemployed Profile	60
1.8	Log Earnings - Cumulative Years OLF Profile	61
1.9	The Effect of Unemployment on Earnings by Timing of Unemployment	62
1.10	The Effect of OLF on Earnings by Timing of OLF periods	63
1.11	High Type Belief	64
2.12	Application Process to Traditional Universities	92
2.13	Smoothed Language PAA Score Distribution	92
2.14	Smoothed Math PAA Score Distribution	93

2.15	Graduation from Selective University Discontinuity	94
2.16	Discontinuity at Pre-treatment Outcomes	95
2.17	Earnings Discontinuity by Experience	96
3.18	Major Trends (1990 to 2012)	139
3.19	H-1 and H-1B Visa Population	140
3.20	Model and Counterfactual (1/2)	141
3.21	Model and Counterfactual (2/2)	142

CHAPTER I

Schooling, Experience, Career Interruptions, and Earnings

1.1 Introduction

Economists have long recognized that schooling and experience are two of the most important aspects of earnings determination.¹ Given the importance of these two variables, a natural question is: how does their interaction affect earnings? In other words, do educated workers have a higher or lower wage increase as they accumulate experience? Which theories can explain this relationship?

In addition to the time spent at work, it is also well documented that individuals spend a significant portion of their time unemployed or out of the labor force during their careers, and these events have a persistent effect on a worker's life-time earnings.² Given the importance of non-working events throughout a worker's life cycle, it is also of considerable interest to

¹The study of the impact of schooling and experience on earnings goes back to Becker (1962), Mincer (1962) and Ben-Porath (1967).

²Examples of papers studying the effect of career interruptions on earnings include Mincer and Polachek (1974), Corcoran and Duncan (1979), Mincer and Ofek (1982), Kim and Polachek (1994), Light and Ureta (1995), and Albrecht et al. (1999).

investigate whether educated workers suffer greater or lower wage losses after out-of-work periods.

It is within this context that this paper examines how the interaction between schooling and work experience affects earnings. In contrast to the existing literature, I take into consideration that workers spend a significant amount of time not employed throughout their careers and that working and non-working periods are substantially different in terms of the interaction between workers and firms.

While considering the difference between working and non-working periods seems natural, this difference has been ignored in most of the existing literature. Table 1.1 presents some of the most important papers that have addressed how the interaction between schooling and experience affects earnings. As can be seen in the table, in order to identify whether more educated workers have a higher or lower wage increase with experience, these papers have used rough measures of experience, such as age minus schooling minus six or years since transition to the labor force.³

As seen in the table, the overall finding in the literature is that returns to potential experience *do not change* across educational groups in old datasets (Mincer, 1974), or *decrease* with educational level in the most recent datasets (Lemieux, 2006 and Heckman et al., 2006).⁴ These results had a lasting influence on empirical work in the field of labor economics. For example, Mincer (1974) used his findings to justify the separability between schooling and

³Note also that these studies differ on how they define earnings. For example, Farber and Gibbons (1996) use earnings in levels. There is also a difference between using annual or hourly wages as the dependent variable. Mincer (1974) uses annual earnings, but only finds evidence for parallel wage profile when controlling for weeks worked in the past calendar year. With the exception of Heckman et al. (2006), most recent papers have used hourly earnings as the dependent variable.

⁴Altonji and Pierret (2001) also find negative coefficients for interaction between schooling and experience when including ability measures not observed by firms that are correlated to schooling on the earnings equation.

experience present in the Mincer earnings equation, which has remained for decades the “workhorse” of empirical research on earnings determination.⁵

Despite the unquestionable value of the articles presented in table 1.1, in this paper I point out issues associated with the measures of experience they use. The first contribution of this paper is to demonstrate that potential experience used in Mincer (1974) confounds the impact of two distinct events on earnings: actual experience and past non-working periods. Furthermore, I demonstrate that if educated workers suffer greater wage losses after out-of-work periods, *potential experience* can produce greater bias to the returns to experience for more educated workers.

This result is at odds with the literature that discusses the bias associated with using potential experience variable (Filer, 1993, Altonji and Blank, 1999, and Blau and Kahn, 2013). According to this literature, potential experience generates *lower* bias to the returns to experience for demographic groups with higher employment attachment, such as more educated workers. In contrast, I demonstrate that potential experience can generate a *greater* bias to the returns to experience for workers with higher employment attachment if their earnings are more affected after career interruptions. To my knowledge, this is the first paper that addresses this matter.

The second contribution of this paper is to use the National Longitudinal Survey of Youth (NLSY) to estimate a model where earnings depend on work experience, past unemployment and non-participation periods, and their interaction with schooling. While there is an

⁵The existing literature presents some possible explanations for the non-increasing effect of schooling on earnings. In the traditional Mincerian model, all workers have the same rate of returns to on-the-job investment, that is independent of their educational achievement. This independence between human capital investments at school and on-the-job can justify the parallel log earnings-experience profiles across educational groups. On the other hand, Farber and Gibbons (1996), and Altonji and Pierret (2001) claim that schooling is used by employers as a signal of a worker’s ability. Therefore, schooling should not become more important for earnings as a worker accumulates work experience.

extensive literature the analysis the impact of career interruptions on earnings (Mincer and Polachek, 1974 and Mincer and Ofek, 1982), on the gender gap (Kim and Polachek, 1994 and Light and Ureta, 1995), and on race wage gap (Antecol and Bedard, 2004), to my knowledge this is the first paper that addresses how past working and non-working periods affect the wage coefficient on schooling.

The results from the estimation of the earnings model that fully characterize the work history of individuals are remarkably different from the standard specification using potential experience. Using my preferred estimation method, I find that for non-black males, the wage coefficient on schooling *increases* by 1.8 percentage points with 10 years of actual experience, but *decreases* by 2.2 percentage points with 1 year of unemployment. In other words, the earning differential between the more and less educated workers mildly rises with actual experience but significantly falls with unemployment. The periods a worker spends out of the labor force (OLF) do not significantly affect the wage coefficients on schooling. I also find qualitatively similar results for blacks and women, with the exception that I find a negative effect of the interaction between OLF periods and schooling on earnings for women.

I provide several robustness checks for these results. First, I estimate a non-parametric model where I do not impose restrictions on the relation between earnings, schooling, work experience, and career interruptions. Second, I change the earnings model so that the timing of career interruptions can also change the effect of schooling on earnings. Third, I take into consideration the possible endogeneity of work history, and estimate a model using an individual fixed effect assumption. In all these specifications, I consistently find that more educated workers have a higher wage increase with work experience but suffer greater wage losses after unemployment periods.

Given the novelty of these results, my third contribution is to propose a model that can

rationalize the empirical findings of this paper. In the model, the productivity of a worker depends on his ability, schooling level and work experience. However, different from the articles in table 1.1, I assume that ability is complementary to schooling and experience in determining a worker's productivity. In other words, high ability workers have higher returns to human capital investments at school and on-the-job.

The model also shares features with the employer learning literature (Farber and Gibbons, 1996 and Altonji and Pierret, 2001). Firms observe schooling and the work history of individuals but have imperfect information about their ability. Within this framework, employers have to make predictions about a worker's ability using the information available at each period. As low-ability workers are more likely to be unemployed throughout their careers, firms use information regarding past employment history of workers in the prediction of their unobservable ability.⁶

Based on the framework described above, the model can predict the empirical results of the paper. The intuition is as follows. High ability workers learn faster on-the-job and have a higher productivity growth as they accumulate work experience. As ability and education are positively related, the model predicts that educated workers have a higher wage increase with experience. In addition, employers use past unemployment as a signal that a worker is low ability. As low ability workers have lower returns to schooling, the model predicts that more educated workers suffer greater wage losses when their low ability is revealed through unemployment.

The paper is organized as follows. In section 2, I discuss the issues of using potential experience when estimating a typical Mincer equation if career interruptions have an impact

⁶While there are studies where employers use lay off information (Gibbons and Katz, 1991) or the duration of an unemployment spell (Lockwood, 1991 and Kroft et al., 2013) to infer a worker's unobservable quality, in this paper firms take into consideration the full work history of an individual.

on earnings. In section 3, I describe the data and I show some descriptive statistics. Section 4 presents the main empirical results of the paper, and I provide some robustness checks. In section 5, I describe the model, and Section 6 concludes the paper.

1.2 Potential Experience and Career Interruptions in the Mincer Equation

The Mincer earnings equation has long been long used as the workhorse of empirical research on earnings determination. Based on theoretical and empirical arguments, Mincer (1974) proposed a specification where the logarithm of earnings is a linear function of education and a quadratic function of potential experience (age minus schooling minus six). Mincer also suggested that schooling and experience are separable in the earnings equation, meaning no interaction term between these two variables is required in the earnings equation. Notably, as discussed shown in table 1.1, Mincer finds evidence that potential experience profiles are nearly parallel across educational groups.

There is wide discussion on the potential pitfalls with the earnings specification proposed by Mincer (Murphy and Welch (1990) and Heckman et al. (2006)), including discussion of the issues associated with using the potential experience variable (Filer, 1993 and Blau and Kahn, 2013). In addition to the existing critiques, in this section I discuss new issues with using the potential experience variable when non-working periods affect earnings.

In order to give some perspective to this issue, I begin the analysis with the traditional case where earnings are affected only by actual experience and not by non-working periods.⁷ The

⁷In Mincer (1974), the potential experience variable is interpreted as a measure of on-the-job training.

log-earnings generating process of worker i at time period t ($\ln w_{it}$) with level of schooling s is defined by the equation below. The parameter β_1^s identifies the impact of the increase of actual experience for workers with level of education s .

$$\ln w_{it} = \beta_0^s + \beta_1^s \text{exper}_{it} + \varepsilon_{it} \quad (1.1)$$

For expositional purposes, and different from Mincer's suggested specification, I assume that log earnings are a linear function of experience. As discussed in Regan and Oaxaca (2009), an inclusion of a quadratic and cubic term tends to exacerbate the type of bias that is discussed here.⁸

The object of interest of the paper is the interaction between schooling and experience. In terms of the equation above, I am interested in how the parameter β_1^s changes across different educational groups. Note that according to Mincer (1974) original specification, log-experience profiles are parallel across educational groups: $\beta_1^s = \beta_1$ for all s .

Equation (1.1) describes how log-earnings changes with actual experience, but individuals can spend some time not working after they leave school. I define $l_{\tau i}$ as an indicator variable that assumes a value of one if individual i worked at past time period τ . The actual experience variable is defined by the sum of past working periods after a worker left school:

$$\text{exper}_{it} = \sum_{\tau=g}^{t-1} l_{i\tau}$$

where g is the time at which an individual leaves school. I also assume that $l_{i\tau}$ is independent of the wage error term ε_{it} and $\mathbb{E}[l_{i\tau}] = p_s$, where p_s is a constant between zero and one that

⁸In the empirical section of the paper I include different functional forms for both actual experience and career interruptions.

indicates the expected fraction of periods that a worker with s level of education stays employed after leaving school.⁹

In most datasets it is not possible to identify an individual's work history. For this reason, researchers have long used rough measures of experience which do not distinguish working and non-working periods, such as the potential experience variable. In the context described above, the potential experience variable $pexp_{it}$, is defined as the time period since an individual left school:¹⁰

$$pexp_{it} = t - 1 - g$$

In this framework, it is easy to show that the coefficient that expresses how earnings change with potential experience is a biased estimator of β_1^s , such that $\tilde{\beta}_{pex} = p_s \beta_1^s$. In fact, this is typical attenuation bias associate with using the potential experience variable present in the literature (Filer, 1993 and Blau and Kahn, 2013). Note that while potential experience attenuates the returns to experience for all demographic groups, the attenuation bias is higher for demographic groups with lower employment attachment. That is the reason why using complete measures of actual experience is a special issue in the literature that studies the gender wage gap (Altonji and Blank, 1999).

Note that as educated workers tend to have a higher employment attachment than uneducated workers, such that $p_s > p_{s-1}$, this model predicts that potential experience underestimates the difference in the wage growth between educated and uneducated workers. In other words, if earnings are not affected by career interruptions, the potential experience

⁹A discussion of the potential endogeneity of $l_{i\tau}$ is found in section 1.3.2.

¹⁰Mincer (1974)'s definition of age-6-schooling would also generate an error term regarding the correct measure of the time a worker left school. For simplicity, I will assume this term is orthogonal to all other variables of the model, and therefore, I ignore it here.

generates a lower bias to the returns to experience for more educated workers. However, as I will show in section 1.3.2, this is the opposite to what it is observed in the data.

Suppose now that in addition to actual experience, non-working periods also have a long-term impact on wages.¹¹ A representation for the earnings equation in this framework would be:

$$\ln w_{it} = \beta_0^s + \beta_1^s \text{exper}_{it} + \beta_2^s \text{interr}_{it} + \varepsilon_{it} \quad (1.2)$$

where interr_{it} is a measure of career interruptions of a worker since leaving school. Using the same notation as before, I define career interruptions as the accumulation of non-working periods since an individual left school:

$$\text{interr}_{it} = \sum_{\tau=g}^{t-1} (1 - l_{i\tau})$$

Note that for simplicity, I assume that earnings are affected by the cumulative non-working periods. However, one can argue that the order and length of non-working periods have a different impact on earnings (Light and Ureta, 1995). In the empirical sections of the paper I also consider this possibility, but for exposition I assume that earnings are only affected by the accumulation of out-of work periods (Albrecht et al., 1999).

Under the earnings generating process described in (1.2), it is easy to show that a regression of earnings on potential experience identifies the following object:

¹¹Possible explanations for that are human capital depreciation (Mincer and Polachek, 1974 and Mincer and Ofek, 1982), firms using the information on past non-working periods as a signal of a worker's productivity (Albrecht et al., 1999), or even that workers accept a wage loss after career interruptions due to liquidity constraints, end of non-working benefits or disutility from leisure (Arulampalam, 2001). In section 1.4, I present a theory for why career interruptions affect wages.

$$\tilde{\beta}_{pex}^s = p_s \beta_1^s + (1 - p_s) \beta_2^s$$

Note that in this framework $\tilde{\beta}_{pex}^s$ confounds the effect of actual experience and career interruptions on earnings. In precise terms, the potential experience effect on earnings is a weighted average of β_1^s and β_2^s , with the weight being defined as the expected employment attachment of workers.

A few comments are needed on how this framework is related to the traditional potential experience bias when career interruptions do not have an impact on earnings, as presented in the discussion of model (1.1). First, if career interruptions have a negative impact on earnings ($\beta_2^s < 0$), the downward bias on estimating on the returns to actual experience is even greater than what the literature has been suggested (Filer, 1993 and Blau and Kahn, 2013).

Second, the potential experience bias can cause greater bias for groups with higher employment attachment. If demographic groups with high employment attachment are also more affected by career interruption (more negative β_2^s), it might be the case that $\tilde{\beta}_{pex}^s$ is a more biased estimator of β_1^s than it is for groups with low employment attachment. In section 1.3.2 I demonstrate that i) educated workers have a higher employment attachment; ii) educated workers face much greater wage losses with career interruptions; and iii) potential experience produces a greater bias to the returns to actual experience for more educated workers.

1.3 Empirical Dynamics

1.3.1 Data

The data used in this paper are the 1979-2010 waves of the National Longitudinal Survey of Youth (NLSY) 1979. The NLSY is well suited for this study because it contains detailed information about individuals' work history since an early age, and follows them during a significant portion of their careers. The individuals in the sample were 14–22 years old when they were first surveyed in 1979, and they were surveyed annually from 1979 to 1993 and biennially from 1994 to 2010.

The sample is restricted to the 2,657 non-black males from the cross-section (nationally representative) sample. This decision to restrict the sample was based on several reasons. First, this is a more stable demographic group during the decades of analysis. The labor market for women and blacks has passed through significant changes in the past 30 years. Second, reasons for career interruptions might differ by gender and race. Even though it is possible to differentiate unemployment from out-of-the-labor-force periods in the data, it is well known that reasons for non-participation in the labor market can substantially differ among demographic groups. Finally, most of the current studies presented in table 1.1 restrict the sample to non-black males, consequently this sample restriction allows a better comparison between my results and previous studies. Nevertheless, given the interest in women and blacks, I also present the main results of the paper for these groups, separately. I define “year of leaving school” as the year when a worker has achieved his highest schooling level and I consider only workers that have been in the labor market after they left school.¹²

¹²I dropped 75 individuals that did not have any observations after the year they left school.

Note that this definition for year of leaving school assures that career interruptions are not caused by a worker's decision to go back to school. However, it also ignores work experience that an individual might have accumulated before achieving his highest degree level. In order to show that the main findings of the paper are not sensitive to such previous work experiences, I also present robustness checks where I define "year of leaving school" as the year an individual reports to not be enrolled in school for the first time.

In the NLSY it is possible to identify week-by-week records of individuals' labor force status since 1978. I use these variables to calculate for each potential experience year (age minus schooling minus six) the share of weeks that each worker in the sample spent working, unemployed, out of the labor force, or in active military service. I use this information to present statistics on average employment attachment over the life cycle for high school graduates and workers with at least a college degree in figures 1.1 and 1.2, respectively. These figures reveal that both high school and college graduates spend on average a significant share of their time after leaving school not working, although career interruptions happen much more often for the former group.

A surprising finding from these figures is that non-black males spend a significant share of their time out-of-the labor force throughout their careers. Although the NLSY provides limited information on the reasons for non-participation of workers, I did some further investigation of the available data for why these group of workers are out-of-the labor force.¹³ The results show that the reasons are very diverse, with the three most common reasons being individuals that did not want to work (20%), had a new job they were to start (19%) and were ill or unable to work (14%).

¹³The data is limited due to the changes of questionnaires across years. These statistics are based on the years 1989-1993, when the most complete questionnaires on the reasons for non-participation are available.

In addition to week-by-week information, NLSY also provides information on weeks between interview years that an individual spent working, unemployed, out of the labor force, or in military service¹⁴ These retrospective variables were used to construct the main work history variables used in the paper, as presented in table 1.2. More specifically, for each individual, work experience is defined as the cumulative number of weeks spent working since leaving school. In addition, cumulative unemployment, OLF, and military service years were defined as the number of weeks spent in each of these labor force conditions since leaving school. I then divide all variables by 52, so that the measurement unit is year.¹⁵ Throughout the paper, potential experience is defined as age minus schooling minus six. This is the variable typically used in the literature (table 1.1) to measure experience, and as discussed before, it does not distinguish working and non-working periods throughout a worker's career. Note that because some individuals take more time to finish school than their schooling years, potential experience does not accurately measure the years a worker is in the labor market. For this reason, I also use time since leaving school as an alternative measure of experience that does not account for non-working periods. Note that time since leaving school is just the sum of the other cumulative work history variables.¹⁶

The wage is calculated as the hourly rate of pay (measured in year 1999 dollars) for the cur-

¹⁴There is also information on the percentage of weeks that NLSY cannot be accounted for. I use this information as a control in all regressions.

¹⁵In section 1.3.2 I also explore the possibility that timing of career interruptions might affect earnings.

¹⁶An issue I faced while creating the work history variables is the fact that 7% of the individuals in the sample graduated before 1978 and there is no available information regarding their work history before this year. I try to overcome this problem by using information available on when a worker left school (a year before 1978) and impute the work history variables described in table 1.2 for these individuals, between the year of leaving school and the year 1978. The imputation method consists of calculating the number of work/unemployment/OLF/military service weeks for the 1978 calendar year, and the assumption that it was constant between the year of leaving school and 1978. An alternative approach is to drop the 196 individuals who graduated before 1978 from the analysis. The results of this second approach are quite similar to imputing the work history variable, so I decided to omit them in this paper, but they are available upon request.

rent or most recent job of a worker.¹⁷ In order to perform the earnings equation estimation, I also restrict the observations to individuals employed at time of interview who work for hourly wages higher than \$1 and less \$100.¹⁸ After these sample restrictions given above, the remaining sample consists of 2,484 individuals with 33,707 observations. All the statistics in the paper are unweighted.

Table 1.3 contains the main statistics of the sample used in the earnings equation estimations for different educational levels. This table highlights some important features of the data. First, the mean of the potential experience and time since leaving school variables are significantly greater than the mean of the work experience for all educational groups. This shows that even for non-black males – a group with considerably higher employment attachment – potential experience substantially overstates actual experience. However, as expected, the difference is higher for less educated workers. Second, the individuals in all the educational groups spend more time out of the labor force than unemployed throughout their career. Finally, the work history information reported in the NLSY is quite accurate: for only 0.8% of weeks since leaving school NLSY was not able to define the labor status of the workers in the sample.

1.3.2 Earnings Dynamics Estimation

There are two main earnings models that are estimated in this paper. The first model represents the typical earnings equation that has been widely used in the literature, which

¹⁷The hourly rate of pay is calculated in the NLSY from answers to questions concerning earnings and time units for pay. If a respondent reports wages with an hourly time unit, actual responses are reported as the hourly rate of pay. For those reporting a different time unit, NLSY uses number of hours usually worked per week to calculate an hourly rate of pay.

¹⁸There are 41 individuals who do not have any observations during the whole period of analysis with earnings within this interval.

shows how the effect of schooling on wages changes with potential experience (see table 1.1). I refer to this model as the traditional model and define log-earnings of individual i in time period t as:

$$\ln w_{it} = \alpha_0 + \alpha_1 s_i + \alpha_2 (s_i \times pexp_{it}) + g(pexp_{it}) + \varepsilon_{it} \quad (1.3)$$

where $\ln w_{it}$ is the log of hourly earnings, s_i is years of schooling and $pexp_{it}$ is the potential experience, defined as “age - schooling - six” or “time since graduation”, which do not distinguish working and non-working periods and $g(\cdot)$ is as cubic function.¹⁹ The primary interest of the paper is estimating the parameter α_2 which identifies how the wage coefficient on schooling changes with potential experience. It is important to note that in previous work (table 1.1) this parameter has been consistently estimated as non-positive; I aim to test whether the same result is found in the sample used in this paper.

In addition to equation 1.3, I also estimate a wage model that fully characterizes the past employment and unemployment history of workers:

$$\ln w_{it} = \beta_0 + \beta_1 s_i + \beta_2 (s_i \times exper_{it}) + \beta_3 (s_i \times interr_{it}) + f(exper_{it}) + h(interr_{it}) + u_{it} \quad (1.4)$$

where $exper_{it}$ is work experience and $interr_{it}$ is a measure of career interruptions since leaving school. The objects of interest are the parameters β_2 and β_3 , which identify how the wage coefficient on schooling changes with work experience and past non-working periods

¹⁹Mincer (1974) uses log of annual wages and $g(\cdot)$ function is defined as a quadratic function. But since the seminal paper from Murphy and Welch (1990), the convention is to use log of hourly earnings and define $g(\cdot)$ a cubic (or even quartic) polynomial.

respectively.

When modeling an earnings function that accounts for the work history of individuals, a researcher is confronted with some non-trivial choices. First, there is a question regarding the appropriate way to measure career interruptions. It has been shown that different labor force status of individuals during career interruptions might have different impact on subsequent wages (Mincer and Ofek, 1982 and Albrecht et al., 1999). For this reason, I will follow the literature and make the distinction between periods of unemployment, time spent out of the labor force, and military service periods.

Second, one can claim that the timing of career interruptions is also important for earnings determination. With respect to this issue, the literature has suggested different specifications, ranging from the simple accumulation of out-of-work periods since leaving school (Albrecht et al., 1999) to a less parsimonious model, which characterizes the number of weeks out of employment for every year since leaving school (Light and Ureta, 1995). For the main results of the paper I will follow Albrecht et al. (1999) and accumulate periods of unemployment and out-of-work since leaving school. However, in subsection 1.3.2 the analogous results using a less parsimonious model are also presented, where timing of non-working periods is important for earnings.

The final non-trivial choice is how to define the functions $f(\cdot)$ and $h(\cdot)$. In order to be consistent with the most recent literature on the earnings equation (Murphy and Welch (1990)), I define $f(\cdot)$ as a cubic polynomial in the main tables of the paper. By analogy, I will also define $h(\cdot)$ as cubic polynomial, although the coefficients of higher order terms are usually not significant. Nevertheless, I will also present a less-restricted model, where I estimate both $f(\cdot)$ and $h(\cdot)$ non-parametrically in subsection 1.3.2 and the results are qualitatively similar to the ones presented with the cubic assumption.

Main Results

Throughout the paper I normalize the interactions between schooling and measures of work history variables such that coefficient of interactions represent a change in the wage coefficient on schooling with 10 years of experience, unemployment, or OLF periods. All the standard errors presented are White/Huber standard errors clustered at the individual level.

Columns (1) and (2) of table 1.4 show the estimation of the traditional earnings model as presented in equation (1.3). The main point of these estimations is to show that one can replicate the finding of the literature, as presented in table 1.1, using the sample restrictions of this paper. First, in column (1) I estimate that the effect of an extra year of schooling on earnings in the beginning of a worker career is 11% (0.006). Next, I estimate that interaction between schooling and potential experience is statistically insignificant. This result is in accordance with Mincer (1974), who found no effects of the interactions between schooling and potential experience on earnings (parallel or convergence of log earnings potential experience profiles across educational groups). Finally, in column (2) I estimate the same specification using time since leaving school as a measure of experience. This measure also does not distinguish working in non-working periods but accurately identifies the period in which a worker left school. Note that the results from these specifications are similar to the ones presented in column (1).

Column (3) provides the estimation of the career interruptions earnings model as presented in equation (1.4). As can be seen, the result from this specification is remarkably different from the ones using the traditional model. First, I estimate a lower schooling coefficient of 8% (0.005). Second, I find a positive and significant coefficient of 0.018 for the interaction between schooling and work experience, meaning that the effect of one additional year of

education increases from 8% to 10%, after a worker accumulates *ten* years of work experience. Furthermore, I estimate a negative effect of the interaction between past unemployment and schooling. Specifically, I estimate that the wage coefficient on schooling decreases by 2.1%, following *one* year of unemployment. Finally, I find a positive – but not significant – interaction between OLF periods and schooling.²⁰ But, as discussed in section 1.3.1, the interpretation for the impact of OLF periods on wage for this demographic group is challenging due to heterogeneous reasons that lead to this type of career interruption.

Columns (4) and (5) provide more robustness to the previous results. In column (4) tenure and its interaction with schooling are added to the model. The idea behind this addition is to investigate whether the main findings of the paper are due to the period a worker is attached to a particular employer, rather than general labor market experience. From these estimations, I find that: i) the coefficients of the career interruptions model are barely affected by the inclusion of these variables; and ii) the wage coefficient on schooling is not significantly affected by tenure. This result suggests that firm-specific mechanisms are not the main explanation for the empirical findings of the paper. This is the approach that is followed in section 1.4.

In column (5) Armed Forces Qualification Test score (AFQT) and its interaction with work experience are added to the earnings equation.²¹ The AFQT score has been used in the employer learning literature (Farber and Gibbons, 1996 and Altonji and Pierret, 2001) as a measure of a worker’s ability that is not easily observed by firms. According to this literature, when AFQT is included with its interaction with experience in the earnings equation, it causes the decreasing with experience (as described in table 1.1). Note that this result is not

²⁰I also reject with 99% confidence that the coefficient of the interaction between schooling and unemployment is equal to the coefficient of the interaction between OLF periods and schooling.

²¹AFQT is standardized by the age of the individual at the time of the test.

found in a model that accounts for career interruptions of workers: while there is a decline of β_2 from columns (3) to (5), the coefficient is still positive and significant. In addition, the other coefficients of interest remain practically unchanged with the inclusion of AFQT in the equation.

Figures 1.3, 1.4, and 1.5 illustrate how the wage coefficient on schooling changes with the work history variables used in the paper. In these figures I report the coefficients of schooling with a 95% confidence interval estimated from the same earnings model as presented in column (3) of table 1.4. The only difference is restricting the sample to workers within a specific range of work history variable (as presented in the x-axis) and the omission of the interaction terms between schooling and work history variables from the equations.

Based on this approach, figure 1.3 shows a wage coefficient on schooling of 8% for workers with 0 to 4 years of work experience. However, this coefficient rises for workers with higher experience levels. In precise terms, I estimate the effect of schooling on earnings at 11% for workers with 16 to 20 years of work experience. In contrast, figure 1.4 shows that the wage coefficient on schooling tends to decrease for workers with higher levels of cumulative unemployment. In fact, I estimate that workers with 0 to 0.4 cumulative years of unemployment have a 10% wage coefficient on schooling, while workers with cumulative years of unemployment between 1.6 and 2 are rewarded only 4% for an extra year of education. Finally, figure 1.5 shows that the wage coefficient on schooling does not change significantly within OLF groups. All these results are consistent with the findings of table 1.4.

As discussed in section 1.3.1, the main group of interest for this work is non-black males. Nevertheless, one might be interested on the empirical results for other demographic groups. In table 1.5, I present the results of the career interruption model for black males, non-black females and black females in columns (1), (2) and (3) respectively. The main findings are

similar to those for non-black males. For black males and non-black females, I estimate: i) a positive and significant effect of the interaction between work experience and schooling; and ii) a negative effect of the interaction between past unemployment and schooling on earnings. Neither work experience nor cumulative unemployment have a significant effect on the returns to schooling for black females. Finally, past OLF periods have a negative impact on the returns to schooling for both non-black and black females. However, it is well-known that reasons for non-participation periods are substantially different for males and females, which poses a challenge for comparing the results for these two groups.

Finally, table 1.6 provides robustness check that the main results of the paper are not sensitive to the definition of the year of leaving school. In precise terms, and different from the other results of the paper, in this table a worker enters the labor market when he first leaves school and the accumulation of work, unemployed and OLF weeks start in this period. As discussed before, on one hand, some of the career interruptions can be justified by a decision of a worker to return to school after spending some time in the labor market. On the other hand, I can account for employment periods a worker had before returning to school in the construction of the work experience.

The table shows that the results using this definition for year of leaving school is very similar to the ones presented in table 1.4. In fact, in column (1) I estimate a 7% effect of schooling on earnings at the beginning of a workers career. Second, there is a positive and significant coefficient of interaction between schooling and work experience of 0.018. In contrast, there is a negative effect of the interaction between past unemployment and schooling of 0.151 and insignificant effect of OLF periods on the returns to schooling. In addition, in columns (2) and (3) I find similar results when including tenure and AFQT and its interactions with schooling and work experience respectively on the wage equation.

Earnings Profiles and Nonparametric Regressions

In this subsection I estimate a less restricted earnings model without imposing functional form assumptions on the relation between work experience, cumulative unemployment, and OLF years and earnings. In these estimations I also substitute years of schooling with educational degree dummies. This procedure allows the model to account for non-linearity in the relation between schooling and earnings. The earnings profiles are plotted with respect to work experience, cumulative years unemployed, and cumulative years OLF for different educational groups. The estimated non-parametric model is the following:

$$\ln w_{it} = f_s(\text{exper}_{it}) + h_s(\text{cunemp}_{it}) + g_s(\text{colf}_{it}) + \eta_{it} \quad (1.5)$$

where s represents educational group variables: less than high school, high school degree, some college and bachelor degree or more. As before exper_{it} is work experience. I also define cunemp_{it} as the cumulative years a work spent unemployed, and colf_{it} as the cumulative years a worker spent OLF. Different from model (1.4), there is no imposition of any parametric restriction on $f_s(\cdot)$, $h_s(\cdot)$ and $g_s(\cdot)$. However, I still impose the additive separability of the work history variables in the model. The method used for the non-parametric estimation is the differentiating procedure described in Yatchew (1998).²² I use locally weighted regressions using a standard tricube weighting function and a bandwidth of 0.5 when estimating f_s and 0.25 when estimating h_s and g_s .²³

Figure 1.6 plots the estimate of $f_s(\cdot)$ for different educational groups. The figure shows

²²In this method, I estimate each function $f_s(\cdot)$, $h_s(\cdot)$, $g_s(\cdot)$ separately, imposing a functional form assumption for the non-estimated functions. In precise terms, when estimating $\hat{g}_s(\cdot)$, I assume that $f_s(\cdot)$ and $h_s(\cdot)$ are cubic polynomial but impose no parametric restriction on $g_s(\cdot)$. The same procedure is applied when estimating $\hat{f}_s(\cdot)$ and $\hat{h}_s(\cdot)$.

²³The overall results of this graph are not sensitive to the choice of different bandwidths.

that the log earnings-work experience profiles have a concave shape as previously found in the literature (Murphy and Welch, 1992), with wages growing faster at the beginning of a worker's career. In contrast to previous literature, I estimate a much steeper wage growth for more educated workers, than for uneducated workers. In fact, the figure shows that the wage gap between individuals with at least a college degree and other workers tends to increase as workers accumulate actual experience. Similarly, the wage gap between high school graduates and workers with less than a high school education is smaller than it is for workers with zero work experience, but increases significantly as workers accumulate experience. These results are in accordance with the findings presented in table 1.4, namely that the wage coefficient on earnings increases, as workers accumulate actual experience throughout their careers.

Figure 1.7 presents the non-parametric estimation of the relation between log earnings and cumulative years of unemployment, defined by the function $h_s(\cdot)$ in equation (1.5), for different educational groups. The figure shows that both college and high school graduates are negatively affected by unemployment periods, as wages decline with the accumulation of this variable. However, the rate of wage decline is substantively different across educational groups since workers with a bachelor's degree have a greater wage decline with unemployment. It is also notable that the wages of workers with less than a high school degree are not significantly affected by unemployment.

Finally, figure 1.8 plots the analogous estimation of the relation between log earnings and cumulative years that a worker spends out of the labor force, as described by the function $g_s(\cdot)$. The evidence shows that this relation is quite heterogeneous among the groups. While the earnings of workers with at least a college degree are almost not affected at all by the accumulation of OLF, workers with less than a high school degree face a substantial wage decrease with OLF periods. The interpretation of these results is difficult because

non-participation periods have heterogeneous justifications among workers.

Timing of Career Interruptions

This section addresses whether accounting for timing of career interruptions in the earnings equation can affect the main findings of the paper. For this reason, instead of assuming that wages are affected by the cumulative unemployment and out-of-the-labor-force periods, I estimate the following log wage model separately by educational groups:

$$\ln w_{it} = \beta_0^S + \beta_1^S + f_s(\text{exper}_{it}) + \sum_{j=1}^5 \gamma_j^s \text{unemp}_{it-j} + \sum_{j=1}^5 \alpha_j^s \text{olf}_{it-j} + \eta_{it} \quad (1.6)$$

where s represents educational group variables: less than high school, high school degree, some college, and bachelor degree or more; unemp_{it-j} is the number of weeks a worker spent unemployed in the calendar year that was j years before the interview and olf_{it-j} is the number of weeks a worker spent out of the labor force in the calendar year that was j years before the interview date. For example, for $t = 1993$, the variable unemp_{it-3} reports the number of weeks a worker spent unemployed in 1990 and olf_{it-3} the number of weeks a worker spent OLF in 1990.²⁴ I divide unemp_{it-j} and olf_{it-j} by 52, allowing the coefficients to be interpreted as changes of year units. Finally, I limit the sample to observations of a worker 5 years after leaving school, so past work history variables reflect events that happened after a worker made the transition to the labor market.

Figure 9 plots the estimation of the coefficients γ_j^s with a 95% confidence interval for different s and j . The graph shows a few interesting facts. First, the weeks spent unemployed in the

²⁴These career interruption variables are constructed based on the week-by-week work history information provided by NLSY, which identifies with precision the periods of unemployment and OLF throughout a worker's career.

past calendar year have the highest impact on earnings for all education groups, but the effects are much higher for workers with a bachelor's degree or higher. In precise terms, the estimation shows that spending the previous calendar year unemployed decreased the earnings of this group by 60%. Second, unemployment periods have a long-term impact on earnings, with a significant negative effect of unemployment weeks, which occurred 5 years prior to the interview. While the difference across educational groups is not as strong, this figure shows that educated workers are also more affected by older unemployment periods.

In figure 1.10, the analogous statistics for α_j^s are reported with a 95% confidence interval, showing that periods spent out of the labor force have a negative impact on the earnings of all workers. However, this effect is much lower than those estimated by unemployment periods, and tend to disappear with time. Finally, while it is estimated that college-graduate workers are more affected by past year OLF weeks than educated workers, the differences across educational groups are not as strong for OLF periods as they are for unemployment periods.

Figures 1.9 and 1.10 bring to light how unemployment and OLF periods affect the effect of schooling on earnings. In order to provide a more accurate test regarding whether the returns to schooling change throughout a workers' career – in a model where timing of career interruptions affect wages – I estimate the model below:

$$\begin{aligned} \ln w_{it} = & \beta_0 + \beta_1 s_i + \beta_2 (s_i \times exper_{it}) + f(exper_{it}) + \sum_{j=1}^5 \lambda_j unemp_{it-j} & (1.7) \\ & + \sum_{j=1}^5 \pi_j (s_i \times unemp_{it-j}) + \sum_{j=1}^5 \rho_j olf_{it-j} + \sum_{j=1}^5 \psi_j (s_i \times olf_{it-j}) + \epsilon_{it} \end{aligned}$$

where all the variables have the same definitions as before and s_i is a measure of years of schooling. In this framework, the coefficients of interest are β_2 , which identifies how the

wage coefficient on schooling changes with work experience, π_j which identifies how the wage coefficient on schooling changes with past unemployment periods j years before the interview and ψ_j which identifies how the wage coefficient on schooling changes with past OLF periods j years before the interview.

The result of the estimation of the earnings model 1.7 is presented in table 1.7. While I estimate the model including olf_{it-j} and its interaction with s_i , for the sake of space these coefficients are omitted in the table. The result shows that ψ_j is not significant for any j . As can be seen in the table: first, the wage coefficient on schooling increases with work experience, even in a model where the timing of career interruption matters, as presented in columns (1) - (3). As can be seen, the estimated β_2 is not very different from the one estimated in table 1.4. Second, as column (2) shows, previous unemployment periods have a significant negative impact on earnings, with previous year unemployment having the highest impact. Third, column (3) shows that, although there is an estimated negative effect of all unemployment periods on the wage coefficient on schooling for all years, recent unemployment periods have a higher impact on earnings. The overall interpretation of these findings is that, while timing of unemployment and OLF might matter for earnings determination, this less-restricted model shows similar patterns, in terms of the effect of work experience and career interruptions on the wage coefficient on schooling, as the one presented in subsection 1.3.2.

Individual Fixed-Effects Estimates

An issue that emerged in models that fully characterize an individual's work history is the possible endogeneity problem of actual experience and career interruptions. The main argument is an omitted variable problem. It is possible that there are some variables not observed

in the data that are related to both current wage determination and past employment. For example, workers with higher career aspirations might have higher employment attachment throughout their life-cycle earnings. In both cases, the seriousness of the endogeneity problem depends on how strong the correlation between current and past levels of the earnings residuals is, and whether past residuals are related to the employment attachment of workers.

A popular approach in the literature when dealing with possible endogeneity of work history is based on an individual fixed effect assumption (Corcoran and Duncan, 1979, Kim and Polachek, 1994, Light and Ureta, 1995 and Albrecht et al., 1999).²⁵ The basic idea of this approach is that the factor related to past employment attachment of workers – which causes the correlation of earnings residuals across time – is an individual-specific fixed component. In terms of the model presented in equation 1.4, the fixed effect assumption means that u_{it} can be written as a sum of an individual component ϕ_i and a transitory component η_{it} , both with mean zero and constant variance. While η_{it} is independent of an individual’s work history, the work history variables can be correlated to ϕ_i .

Table 1.8 presents the main results of the estimation of the wage model described by equation (1.4) using an individual fixed effect estimation. Note that as schooling does not change overtime, I cannot identify β_1 when using this estimation strategy. However, it is possible to identify the effect of its interaction with other time-varying variables, such as work expe-

²⁵There are other suggestions in the literature with respect to ways of addressing the possible endogeneity of work history. Mincer and Polachek (1974) suggest using family characteristics, such as education of the partner or number of children, as instruments for previous working and non-working periods of married women. While it is questionable as to how exogenous these variables truly are, there is evidence that family characteristics have a weak relation to employment attachment of non-black males, the main group of interest of this work. Alternatively, Altonji and Pierret (2001) suggest using potential experience ($pexp_{it}$) as an instrument for actual experience, in a model that earnings are not affected by unemployment periods. However, if career interruptions have impact on wages, the potential experience variable is not a validity instrument for actual experience. In this circumstance, $pexp_{it}$ is not redundant (or ignorable) in the log wage expectation, such that: $\mathbb{E}[\ln w_{it}|exper_{it}] \neq \mathbb{E}[\ln w_{it}|exper_{it}, pexp_{it}] = \mathbb{E}[\ln w_{it}|exper_{it}, interr_{it}]$.

rience, tenure, and cumulative years OLF and unemployment. In order to make these new results comparable to the least square estimation, the same specifications are followed in this table as the one presented by the least square estimation of table 1.4.

The overall results from table 1.8 are qualitatively and quantitatively similar to those estimated by the least square estimation of table 1.4. Namely, the wage coefficient on schooling increases significantly as a worker accumulates work experience, and decreases as a worker accumulates unemployment periods. If anything, the fixed effect estimation shows a lower negative coefficient for the effect of unemployment on the returns to schooling. In other words, this new estimation leaves the conclusions based on the OLS regressions intact.

This result is not surprising in light of the findings of existing literature. Mincer and Polachek (1974), Blackburn and Neumark (1995), and Albrecht et al. (1999) have found that coefficients of the earnings model stay virtually unchanged when dealing with the possible endogeneity problem of work history variables. From these results, one can conclude that the endogeneity of work history appears to be less of a problem when estimating career interruptions models.

1.4 Model

The dynamics estimated thus far are puzzling for conventional models of labor market dynamics. Unlike the past empirical literature, my research finds that more educated workers have a higher increase in earnings with actual experience, while suffering greater earnings losses after unemployment periods. This raises the question as to which economic reasons can explain this relationship. Is it possible to conciliate the existing theories for earnings dynamics with these novel empirical findings? In order to answer these questions, in this

section I present an economic model that can rationalize the empirical findings of this paper.

1.4.1 The model environment

A worker enters the labor market in period 0 and lives for T periods. All firms are identical and the only input used in production is labor. Let y_{it} denote a worker's log-productivity in the t -th period after leaving school.

$$y_{it} = \theta_i g(s_i, x_{it}) \tag{1.8}$$

In this specification, θ_i is the worker's ability, $g(s_i, x_{it})$ is a worker's human capital, which is a function of the worker's schooling level s_i and work experience x_{it} . For exposition, I will omit x and s subscripts henceforth. I assume that both θ_i and $g(s, x)$ are positive, and $\partial g(s, x)/\partial s > 0$, $\partial g(s, x)/\partial x > 0$ and $\partial g(s, x)/\partial x \partial s = 0$. The important assumption is that ability and human capital are complementary in determining the log-productivity, which is captured in the multiplicative specification of (1.8).²⁶ An interpretation of the complementary assumption is that high ability workers can more effectively use their human capital at work and therefore have higher returns to schooling and experience.²⁷

Furthermore, there are only two types of workers: high ability θ_H or low ability θ_L . While schooling and work experience are observed, ability is not observed by either employers or workers. All agents have to make their predictions about a worker's ability based on the information available at each period.

²⁶Note that this assumption makes the model different from the studies presented in table 1.1.

²⁷Papers making similar assumptions include: Acemoglu and Pischke (1998), Gibbons and Waldman (2006), and DeVaro and Waldman (2012).

Information structure

The only available information regarding ability in period 0 is a worker's schooling level s . I define p_s as the fraction of workers with schooling level s that are high ability. I assume that p_s is different from zero and one, and it is strictly increasing with a worker's schooling level, meaning that high ability workers are more likely to get more education. Note that in this version of the model, I do not model schooling decision of workers, but this assumption is consistent with the signaling literature (Spence, 1973) where high ability workers have lower costs to acquire education. Nevertheless, later I sketch how the model could be enriched to allow for the endogeneity of schooling.

In addition to schooling, I assume that in every period some new information about a worker's quality becomes available to all firms.²⁸ This new information can be summarized by the signal \tilde{y}_{it} , which can be a good or bad signal, with high ability workers producing a good signal with probability γ_H and low ability workers producing a good signal with probability γ_L , such that $\gamma_H > \gamma_L$. As in Altonji and Pierret (2001), firms will use information on the worker's signals during the past $x - 1$ *employed* periods to infer a worker's unobservable ability. I define $I_{it} = \{\tilde{y}_{i1}, \dots, \tilde{y}_{ix-1}\}$ as the set of observed past signals.

Different from Altonji and Pierret (2001), an individual can be in one of two possible states at each period of his career: working or not working. Firms can also observe the employment history of an individual, which is characterized by the number of periods an individual was employed $x - 1$ (work experience minus one) and the number of periods a worker was unemployed u since leaving school (career interruptions).²⁹ As will be clarified later, em-

²⁸This information consists on past on-the-job performance, new letters of recommendation, interviews, etc.

²⁹Note that by definition $x + u = t$. Given the perfect linear combination between work experience, career interruption, and time since leaving school, one could define the information set available to firms as two of

ployment history gives extra information about a worker's ability and the timing of working and non-working periods will not be important in the equilibrium of this simplified model.³⁰

Timing and actions

At the beginning of each period the sequence of events and actions are as follows:

1. A fraction δ of individuals are unable to work. These are the workers that are moving for personal reasons or are not able to be matched to any employer.
2. The other fraction $(1 - \delta)$ of workers are able to work and draw a new signal \tilde{y}_{it} for the period.
3. The employers make job offers based on information available in the period and the new signal \tilde{y}_{it} .
4. A worker can either:
 - Choose to work in the period. In this case, a worker accumulates one period of work experience and keeps the signal for future wage offers.
 - Choose to not work in the period. In this case a worker accumulates one period of unemployment, while discarding the signal that will not be used for future wage offers.

any of the three variables. For expositional purposes, I choose to present it as work experience and career interruptions. I also ignore the difference between unemployment and out-of-the-labor-force periods.

³⁰This mechanism is consistent with papers where employers use lay off information (Gibbons and Katz, 1991) or the duration of an unemployment spell (Lockwood, 1991 and Kroft et al., 2013) to infer a worker's unobservable quality. However, in this paper firms take into consideration the full work history of an individual.

Note that in the model unemployment can be involuntary or voluntary. Involuntary unemployment is caused by a worker who could not be matched to any employer in a given period (fraction δ), while voluntary unemployment results from a worker's decision to reject any job offer. I assume that firms cannot distinguish between these two types of career interruptions when making future wage offers. The idea is that (low performance) workers can always tell the employers that they did not work in a period because exogenous reasons were preventing them from working. Nevertheless, firms pay close attention to the accumulation of career interruptions, and workers are unlikely be able to justify the long periods of unemployment as involuntary.

Firms' decision

Firms do not discount the future and long term contracts are not allowed. As in Farber and Gibbons (1996) and Altonji and Pierret (2001), I assume that there is free entry of firms and all employers share the same information about a worker's productivity. As a consequence from competition among employers, the wage offered to a worker i in period t is equal to the expected productivity given the information available at the period and the new signal \tilde{y}_{it} .³¹

$$W_{it} = \mathbb{E}[\exp^{y_{it}} | x, u, s, I_{it}, \tilde{y}_{it}] \quad (1.9)$$

An alternative representation of the wage set up is to define $\mu(s, x, u, I_{it}, \tilde{y}_{it})$ as the employers' belief that a worker is high-type based on the information available up to that point. In this framework, I use equation (1.8) to show that the wage level of a worker in period t can be represented by:

³¹Note that the information that a worker chose to work in period t is implicit in the term \tilde{y}_{it} .

$$W_{it} = \mu(s, x, u, s, I_{it}, \tilde{y}_{it}) \exp^{g(s,x)\theta_H} + [1 - \mu(s, x, u, I_{it}, \tilde{y}_{it})] \exp^{g(s,x)\theta_L} \quad (1.10)$$

The wage process presented in equation 1.10 shows the two different roles of work experience x in the model. On one hand, the term $g(s, x)$ represents the productivity increase of a worker as he accumulates work experience. This mechanism is defined as the human capital effect of working on earnings. On the other hand, accumulating employment periods also provides information about a worker's type, which is represented by the term $\mu(s, x, u, I_{it}, \tilde{y}_{it})$. This mechanism is referred to as the information effect of working on earnings. Furthermore, firms will also use information regarding career interruptions u in the assessment of a worker's type.

Worker's decision

I assume that workers are risk neutral and discount the future using a discount rate $\beta > 0$. At each period a worker has access to the same information as firms.³² In this framework, for individuals that are not exogenously unable to work, the work decision in the first $T - 1$ periods of their career is defined by the following Bellman equation:³³

$$V(s, x, u, I_{it}, \tilde{y}_{it}) = \max\{W_{it}(s, x, u, I_{it}, \tilde{y}_{it}) + \beta(1 - \delta)\mathbb{E}[V(s, x + 1, u, I_{it}, \tilde{y}_{it}, \tilde{y}_{t+1})], \\ b + \beta(1 - \delta)\mathbb{E}[V(s, x, u + 1, I_{it}, \tilde{y}_{t+1})]\} \quad (1.11)$$

³²As it would be clear in equilibrium, even if workers have better information regarding their own ability than firms, this information will be irrelevant for their working decisions.

³³In period T , workers make the same decision but do not consider the future.

where b is the utility flow for not-working. This Bellman equation highlights a trade-off associated with the employment decision.³⁴ On one hand, an individual can choose to work, be paid, and accumulate one year of experience. In this case, the signal \tilde{y}_{it} is used for current and future wages offers. On the other hand, a worker could discard the signal, receive non-working benefits and accumulate one period of unemployment. In this case, firms will not be able to distinguish whether the unemployment period was due to a worker's choice or to an exogenous reason. Nevertheless, firms will use the extra non-working period information to update their beliefs about a worker's ability, and this unemployment information will be used for future wage offers.

1.4.2 Equilibrium

Equilibrium is characterized by a function of the state variables $S_{it} = \{s, x, u, I_{it}\}$ and signal \tilde{y}_{it} to the firms' belief that a workers is high type μ_{it} , a wage offer W_{it} and an individual's decision to work in period t . From this general framework, it is possible to derive some predictions of an individual's optimal working strategy and how firms use past employment and unemployment information to update their beliefs about a worker's type.

Proposition 1: For a given state S_{it} , if it is an optimal strategy for an individual to choose to work after a bad signal draw, it is also an optimal strategy to work in case of a good ability draw.

The justification is straightforward: the firms' belief that a worker is high type is greater after a good signal revelation than after a bad signal. As a result, present and future wage

³⁴Note that individuals are exogenously unable to work in period $t + 1$ with probability δ . As the utility from this state is independent of previous work choices, this possibility should not affect an individual's decision to work in period t . In precise terms, the future expected utility of being exogenously unemployed is additive in both terms of the Bellman equation, and therefore is canceled out.

offers must be higher after a good signal than after a bad signal. For this reason, for any given state, a worker is better off taking the job after a good signal draw than he would be working after a bad signal draw.

This proposition has implications for the adverse selection and employer learning mechanism proposed by the model. Firms realize that workers with bad signals are more likely to be unemployed and workers with good signals are more likely to be employed. Even though firms cannot observe signals produced in the non-working periods, or ascertain whether unemployment was caused by an exogenous reason, they use information on career interruptions and employment periods to update their beliefs about a worker's ability.

Separating Equilibrium

The analysis is now restricted to a separating equilibrium where for any given state, individuals always choose to work after observing a good signal draw, and always decide to not work after observing a bad signal draw. This extreme case highlights the mechanisms of adverse selection and employer learning through work history that I want to stress with the model. It also simplifies the calculation of firms' beliefs and wage offers, and the derivations of the predictions of the model.

Some extra assumptions are required in order to guarantee the existence of such a separating equilibrium. First, I assume that high-ability workers always produce a good signal, such that $\gamma_H = 1$, while low-ability workers can produce both good and bad signals: $0 < \gamma_L < 1$. The direct implication of this assumption is that the decision to work after a bad signal is sufficient to reveal to employers that a worker is low type for the rest of his career.

Nevertheless, it might be optimal for an individual to work after it is revealed that he is low

ability. For this reason, I assume the productivity of a low ability worker is always lower than his non-working utility, such that $\exp^{g(H,T)\theta_L} < b$, where H is the highest schooling level a worker can achieve and therefore $g(H, T)$ is the highest human capital level a worker can possibly have. An interpretation of this assumption is that low-ability jobs are so much less rewarding, that workers would never reveal to firms that they are low ability.

Finally, for any state S_{it} , it must be optimal for an individual to choose to work after a good signal. For this reason I impose the following restriction on θ_H :

$$\tilde{\mu} \exp^{g(0,0)\theta_H} + (1 - \tilde{\mu}) \exp^{g(0,0)\theta_L} > b \quad (1.12)$$

where $g(0, 0)$ is the lowest human capital level an individual can possibly have (zero schooling and zero actual experience) and $\tilde{\mu}$ represents the lowest believe a firm can have that a worker is high type in this separating equilibrium. This term is a function of the parameters p_s , δ , and γ_L and T , and is derived in the appendix of the paper. An interpretation of this assumption is that high-performance jobs are very rewarding and an individual would always work after a good signal.

Under these assumptions, for any state S_{it} the optimal choice of an individual, that is not exogenously unable to work, is to be employed if \tilde{y}_{it} is a good signal and to be unemployed if it is a bad signal. Note that the set of signals I_{it} becomes trivial, since individuals only work in good signal periods. In this case, the set I_{it} is equivalent to the employment periods x and therefore will be omitted henceforth.

Within this framework, it is easy to derive the fraction of workers that are employed and unemployed at each period, and the firms' equilibrium belief that a worker is high type. First, the fraction of high-ability individuals that are employed in each period is equal to

the fraction of workers that were not exogenously unable to work: $\mathbb{P}(Work_{it}|\theta_H) = 1 - \delta$. In other words, because high ability workers always draw good signals, the only reason for this type of worker to be unemployed is being exogenously unable to work, which happens with probability δ .

In contrast, low ability workers can be unemployed due to both exogenous reasons or a bad signal draw. Therefore the probability that a low ability individual is working in a period is $\mathbb{P}(Work_{it}|\theta_L) = \gamma_L(1 - \delta)$. For $\gamma_L < 1$, high type worker are more likely to be employed than low type workers at any point of their career.

Note that working events are independent across time. As a consequence, the probability that a worker has x employment periods and u career interruptions conditional on high and low ability level respectively is characterized by a binomial probability function:

$$\mathbb{P}(X = x, U = u|\theta_H) = \binom{x+u}{x} (1-\delta)^x \delta^u \quad (1.13)$$

$$\mathbb{P}(X = x, U = u|\theta_L) = \binom{x+u}{x} (\gamma_L(1-\delta))^x (1-\gamma_L(1-\delta))^u \quad (1.14)$$

where $\binom{x+u}{x}$ is the binomial coefficient of x and $x+u$.³⁵

In this framework, it is simple to characterize how employers learn about a worker's type throughout his career. The prior about a worker's type when he enters the labor market p_s is defined by the fraction of workers with education level s that are high type. However, as a worker progresses in his career, firms use information on his working and non-working

³⁵Note, in this separating equilibrium, the timing of career interruptions is not important for the probability function. Therefore, firms will only use the cumulative values of work history variables to update their beliefs about a worker's type.

periods to update their beliefs. Based on equations (1.13) and (1.14), I use the Bayesian rule to derive the firms' belief that a worker with experience x and career interruptions u is high type:

$$\mu^*(s, x, u) = \frac{\delta^u p_s}{\delta^u p_s + \gamma_L^x (1 - \gamma_L (1 - \delta))^u (1 - p_s)} \quad (1.15)$$

Equation (1.15) presents features regarding how the firms' belief evolves as a worker progresses in his career. At the beginning of a worker's career firms have no information about a worker's history ($x = 0$ and $u = 0$). In this case, the belief that a worker is high type is defined solely by his schooling level, summarized by the prior p_s . As workers progress in their careers, and low ability workers are more likely to be unemployed, firms update their beliefs using information on x and u .

Proposition 2: $\mu^*(s, x, u)$ strictly increases with x and s and strictly decreases with u . Furthermore, $\lim_{x \rightarrow \infty} \mu^*(s, x, u) = 1$ and $\lim_{u \rightarrow \infty} \mu^*(s, x, u) = 0$.

The idea of proposition 2 is simple. The adverse selection mechanism implies that firms use periods of past employment as a good signal of a worker's type, and past unemployment periods as a bad signal of a worker's type. For this reason, a firm's expectation that a worker is high type must increase with x and decrease with u . For T is large enough, firms should be able to recover a worker's type just by updating their beliefs, using previous working and non-working information.

The figure 1.11 illustrates how the work history of an individual affects firms' beliefs about his type. The blue line shows how the belief that a worker with zero career interruptions ($u = 0$) changes as he accumulates work experience x . The red line shows the analogous relation for a worker that suffers one period of career interruption ($u = 1$) in the beginning of his career,

and is the green line of a worker that suffers with two periods of career interruptions ($u = 2$) in the beginning of his career. Note that the lines terminate at the end of a worker's career. The graph shows that firms have a belief p_s that a worker with zero unemployment is high type at the beginning of his career (blue line). As the worker accumulates work experience, this expectation raises to the point where firms are almost certain he is high type at the end of period T . In contrast, a worker who starts to accumulate work experience with one period of unemployment initiates the process from a lower level of expectation than p_s . However, as he gains work experience, firms use the new employment periods to update their beliefs, and the expectation about his type rises. Eventually, the employment information overcomes the signal of one period of career interruptions and $\mu^*(s, x, u)$ catches up with the expectation from a worker with no career interruptions. Finally, the green line shows that a worker with 2 periods of unemployment starts his career from a very low belief level. Although the worker is able to improve the firms' expectations as he accumulates employment periods, the new information is not such magnitude as to overcome the bad signal of two unemployment periods. As a result, the expectation that he is high type never catches up with that of workers with no career interruptions.

Wage determination

Having characterized how firms form and update their beliefs about a worker's ability in a separating equilibrium, I now turn to demonstrating that this learning process has important implications for wage setting. Using the firms' equilibrium belief that a worker is high type - derived in the past subsection - and the wage setting described by (1.9), one can write the equilibrium wage of a worker with schooling level s , work experience x and career interruptions u as follows:

$$W_{it}^* = \mu^*(s, x, u) \exp^{g(x,s)\theta_H} + [1 - \mu^*(s, x, u)] \exp^{g(x,s)\theta_L} \quad (1.16)$$

where $\mu^*(s, x, u)$ is the equilibrium belief that a worker is high type as defined in equation (1.15). From the equation above, it is easy to show that equilibrium wage levels are strictly increasing with respect to schooling and work experience and strictly decreasing with respect to unemployment period. Nevertheless, in this paper we are interested in how the interaction between schooling, work experience and past unemployment periods affects log earnings.

Proposition 3: Under the assumptions of the above model,

$$\frac{\partial^2 \ln W_{it}^*}{\partial x \partial s} > 0 \text{ for any } s, x \text{ and } u.$$

The proof of proposition 3 is presented in the appendix of the paper but the intuition follows from the assumption that ability and human capital are complementary in determining the log-productivity of a worker. More precisely, in the model, the work experience affects earnings in two ways. First, it increases a worker's log-productivity: workers learn more on the job and therefore become more productive as they accumulate x . The complementarity between ability and human capital implies that high ability workers have a higher log-productivity increase with work experience. Note that by assumption, the fraction of workers that are high ability increases with their schooling level. As a consequence, the model predicts that more educated workers have higher returns to experience.

The second way that work experience affects earnings is through a signaling effect. As described in the model, high ability workers are more likely to be employed in the course of their careers. As a consequence, the probability that a worker is high ability increases with his past employment periods. The complementarity between ability and human capital

also implies that high ability workers have higher returns to schooling. Consequently, the model predicts that workers with high levels of work experience also have higher returns to schooling. To sum up, both human capital and signaling mechanisms imply that the interaction between schooling and work experience have a positive effect on log-earnings.

Proposition 4: Under the assumptions of the above model,

$$\frac{\partial^2 \ln W_{it}^*}{\partial u \partial s} < 0 \text{ for any } s, x \text{ and } u.$$

The proof of the proposition is also presented in the appendix of the paper, but the intuition is similar to the one just described. In the model, low ability workers are more likely to be unemployed throughout their careers. These workers are more likely to draw bad signals and therefore more likely to reject low wage offers. As a consequence, the fraction of workers that are low ability increases with u . Note that due to the complementarity between ability and human capital, low ability workers have lower returns to schooling. Consequently, the model predicts that educated workers are those who suffer the most when they have their low ability type revealed with unemployment. Put differently, the interaction between schooling and past unemployment periods have a negative effect on log-earnings.

1.4.3 Some Extensions

In this subsection I discuss the general intuition of two possible extensions of the basic model presented so far.

Unemployment and Out of the Labor Force

A simplifying assumption used so far is that there are only two possible employment status: working and not working. However, in the empirical part of the paper I also distinguish the impact of unemployment and out-of the labor force periods on earnings. In fact, I find that: i) unemployment periods have a higher negative impact on earnings than out-of-the labor force periods; and ii) there is no significant difference across educational levels in terms of wage losses after OLF periods. A natural question is: how can one incorporate the distinction between unemployment and OLF into the model?

Note that in the model, non-working periods can be explained by a fraction δ of individuals that are unable to work due to exogenous reasons or by individuals that did not work after receiving a bad signal. As described before, firms cannot distinguish between these two types of career interruptions when making future wage offers. A simple way to incorporate OLF periods to the model is to assume that firms can identify a share of non-working periods caused by the exogenous reasons. For example, one can think that workers who were moving due to family reasons can demonstrate to potential employers that they did not work in a period because they were moving. As a consequence, an individual's work history would be also be characterized by the accumulation of non-working periods which are uncorrelated to worker's ability.

Schooling choices

In this paper employment and wage decisions happen after an individual leaves school. Nevertheless, I could also assume that workers have some knowledge of their own innate abilities and make schooling choices at the beginning of their career. In this case, schooling

would also be an endogenous variable of the model. A simple way to introduce schooling decisions is to assume that individuals make education choices in period zero in order to maximize their expected life time earnings, defined by the value function (1.11) in period zero. Even if costs of getting more education do not vary with ability, we would expect that high ability individuals are more likely to achieve higher levels of education because these workers have higher returns to schooling. In this case, ability and schooling would be positively related, such as presented in the model.

1.5 Conclusion

In this paper I extensively examined whether educated workers have a higher or lower wage increase throughout their careers. Different from past work, I accounted for the fact that workers spend a significant amount of time not employed throughout their careers. This distinction is important because, as demonstrated above, the potential experience typically used in previous literature confounds the impact of two distinct events on the earnings: actual experience and past non-working periods. Not surprisingly, I found that these two events have different effects on wages across educational groups. I found that educated workers have a higher wage increase with experience but suffer a greater wage loss after unemployment periods. These results are robust to different specifications of the earnings equation, timing of the unemployment spells, and estimation methods.

In addition, I proposed a model that can rationalize the novel empirical results of this paper. In the model, the productivity of a worker is defined by his observed human capital and his unobserved ability. Firms update their predictions that a worker is high ability as new information becomes available throughout a worker's career. The innovation of the model

is that firms can use past employment and unemployment periods in their assessment of a worker's ability. Under the assumption that human capital and ability are complementary in the determination of a worker's productivity, the model predicts that educated workers have a higher wage increase with work experience but suffer a greater wage loss after career interruptions.

1.6 Theory Appendix

1.6.1 Derivation of $\mu^*(x, u)$ and $\tilde{\mu}$

In the model, firms use Bayesian rule to update their beliefs that a worker is high type based on the past employment history and schooling. For this reason we have that equilibrium belief that a worker is high type given this information is defined as:

$$\mathbb{P}(\theta_H | X = x, U = u, s) = \frac{\mathbb{P}(X=x, U=u | \theta_H, s) \mathbb{P}(\theta_i = \theta_H | s)}{\mathbb{P}(X=x, U=u | \theta_H, s) \mathbb{P}(\theta_i = \theta_H | s) + \mathbb{P}(X=x, U=u | s, \theta_L) \mathbb{P}(\theta_i = \theta_L | s)}$$

Substituting the work history probabilities presented in equations 1.13 and 1.14 and using the prior p_s that a worker is high type, one gets the equilibrium belief that a worker is high type given his work history and schooling level:

$$\mu^*(x, u) = \frac{\binom{x+u}{x} (1-\delta)^x \delta^u p_s}{\binom{x+u}{x} (1-\delta)^x \delta^u p_s + \binom{x+u}{x} (\gamma_L (1-\delta))^x (1-\gamma_L (1-\delta))^u (1-p_s)}$$

dividing both the numerator and denominator by $\binom{x+u}{x} (1-\delta)^x$, I obtain the expression for the equilibrium presented in equation 1.15.

Based on this expression, it easy to see that the lowest possible belief that a worker is high type in a separating equilibrium is defined by workers that were unemployed during T

periods and has zero schooling. Substituting these work history ($s = 0$, $x = 0$ and $u = T$) in equation 1.15, I obtain an expression for $\tilde{\mu}$:

$$\tilde{\mu} = \frac{\delta^T p_0}{\delta^T p_0 + (1 - \gamma_L (1 - \delta))^T (1 - p_0)}$$

1.6.2 Proof of Proposition 3 and 4

The derivations of proposition 3 and 4 come directly from the wage equation presented in equation 1.16. Let's define μ_u^* as the derivative of the equilibrium belief with respect to u , μ_x^* is the derivative with respect to x , and μ_s^* is the derivative with respect to s . From equation 1.15 and assumptions of the model, it is easy to show that $\mu_x^* > 0$ and $\mu_s^* > 0$ and $\mu_u^* < 0$ for every x , s , and u . In other words, the belief that a high type worker is strictly increasing with schooling level and past work experience and strictly decreasing with unemployment.

Using the expressions above and after some tedious algebra, one can show that:

$$\frac{\partial \ln W_{it}}{\partial s \partial x} = \frac{\exp^{g(x,s)(\theta_H + \theta_L)} (\theta_H - \theta_L)}{W_{it}^2} \{ \mu_x^* g_s + \mu_s^* g_x + g_s g_x (\theta_H - \theta_L) \mu^* (1 - \mu^*) \}$$

g_x is the derivative of the human capital function with respect to x and g_s is the derivative of the human capital function with respect to s . Note that by assumption of the model, the human capital function is strictly increasing schooling and experience, so that $g_s > 0$ and $g_x > 0$. Therefore, one can conclude that $\frac{\partial \ln W_{it}}{\partial s \partial x} > 0$ for every x , u , and s . Note also that the human capital effect of experience on earnings is represented by the term $\mu_s^* g_x$ and the signaling effect of experience on earnings is given by the term $\mu_x^* g_s$.

In the same way, one can show that:

$$\frac{\partial \ln W_{it}}{\partial s \partial u} = \frac{\exp^{g(x,s)(\theta_H + \theta_L)} (\theta_H - \theta_L)}{W_{it}^2} \mu_u^* g_s$$

As by assumption we have that $g_s > 0$ as it is easy to show from equation 1.15 that $\mu_u^*(x, u) < 0$, we can conclude that $\frac{\partial \ln W_{it}}{\partial s \partial u} < 0$ for every x , u , and s . The term $\mu_u^* g_s$ identifies that educated workers suffer the most when their low ability type is revealed.

Table 1.1: Literature Review

Study	Data	Dependent Variable	Experience Specification	Sample	Main Findings
Mincer (1974)	U.S. Census, 1960	Log Annual Earnings ³⁶	Age-Schooling-6	White, non-farm, non student men up to age 65.	“Experience profiles of log earnings are much more nearly parallel.”
Faber and Gibbons (1996)	NLSY 1979-1991	Hourly Wage (level)	Time since long-term transition to the labor force	Males and females after long-term transition to the labor force.	“The estimated effect of schooling on the level of wages is independent of labor-market experience.”
Altonji and Pierret (2001)	NLSY 1979-1992	Log Hourly Wage	Age-Schooling-6 ³⁷	White or black males with eight or more years of education.	“Wage coefficients on the variables that firms cannot observe and affect workers’ productivity rise with experience while the coefficient on education falls.”
Lemieux (2006)	CPS 1979–1981, 1989–1991, and 1999–2001	Log Hourly Wage	Age-Schooling-6	Men age 16 to 64 with 0 to 40 years of potential experience.	For 1979-1981 the experience profiles are parallel; For 1989-1991 and 1999-2001 the college-high school wage gap declines as a function of experience.
Heckman et al. (2006)	U.S. Census, 1940-1990	Log Annual Earnings	Age-Schooling-6	White and black males.	“The estimated profiles for white males from the 1940–1970 Censuses generally support the parallelism by experience patterns. Log earnings–experience profiles for the 1980–1990 Censuses show convergence for both white and black males.”

³⁸Mincer only finds insignificant effects of the interaction between schooling and experience when controlling for weeks worked in the past year.

³⁹In Panel 2 of Table 1, the authors present their results using actual experience instrumented by potential experience. I discuss the validity of this approach in section 3.2.4.

Table 1.2: Work History Variables

Variable	Definition
Potential Experience	Age - Schooling - 6
Time since leaving school	Weeks since leaving school /52
Work Experience	Weeks worked since leaving school /52
Cumulative Years OLF	Weeks OLF since leaving school /52
Cumulative Years Unemployed	Weeks Unemployed since leaving school /52
Cumulative Years in Military Services	Weeks in the Military Services since leaving school /52

Table 1.3: Descriptive Statistics

Variable	Education Level			
	Less than High School	High School Degree	Some College	BA or More
Log Hourly Wage (1999 dollars)	2.00	2.25	2.44	2.80
Potential Experience	16.07	14.48	14.91	13.24
Time since graduation	14.21	14.14	12.46	11.52
Work Experience	11.04	11.87	10.86	10.68
Cumulative Years OLF	1.58	1.03	0.82	0.43
Cumulative Years Unemployed	1.41	0.76	0.40	0.24
Cumulative Years in Military Services	0.01	0.36	0.27	0.10
Cumulative Years Unaccounted	0.17	0.12	0.10	0.07
Individuals	224	1,083	508	669
Observations	3,432	16,750	6,138	7,387

Note: See Table 1.2 for definitions of the work history variables.

Table 1.4: The Effect of Schooling, Experience, and Career Interruptions on Earnings

NLSY 1979 - Non-Black Males

Dependent Variable: Log Real Hourly Wage

Method: Least Squares

Model	(1)	(2)	(3)	(4)	(5)
Schooling	0.111 (0.006)***	0.102 (0.005)***	0.082 (0.005)***	0.083 (0.005)***	0.066 (0.006)***
Schooling * Potential Experience/10	0.004 (0.004)				
Schooling * Time since Leaving School/10		0.007 (0.005)			
Schooling * Work Experience/10			0.018 (0.003)***	0.014 (0.004)***	0.014 (0.004)***
Schooling * Cumulative Years Unemployed /10			-0.218 (0.038)***	-0.207 (0.038)***	-0.245 (0.039)***
Schooling * Cumulative Years OLF /10			0.022 (0.024)	0.022 (0.024)	0.031 (0.025)
Schooling * Tenure Years/10				0.002 (0.006)	
AFQT * Work Experience/10					0.020 (0.008)**
Observations	33,707	33,707	33,707	33,181	32,162
R-squared	0.260	0.264	0.321	0.325	0.338
Tenure	No	No	No	Yes	No
AFQT	No	No	No	No	Yes
Other controls:	Cubic Polynomial of Potential Experience and Year Dummies	Cubic Polynomial of Time since Leaving School and Year Dummies	Cubic Polynomial of Work Experience, Cumulative Years OLF/Unemployment/Military; Uncounted Years; and Years Dummies		

*** p<0.01, ** p<0.05, * p<0.1

White/Huber standard errors clustered at the individual level are reported in parentheses.

Note: AFQT is normalized to have a standard deviation of 1. Difference in the number of observations between models (3) , (4) and (5) is due to 526 observations of individuals with missing tenure and 1,545 observations of individuals with missing AFQT information.

Table 1.5: The Effect of Schooling, Experience, and Career Interruptions on Earnings, Other Demographic Groups

NLSY 1979 - Other Demographic Groups

Dependent Variable: Log Real Hourly Wage

Method: Least Squares

Sample	Black Males	Non-Black Females	Black Females
Schooling	0.106 (0.015)***	0.092 (0.005)***	0.111 (0.012)***
Schooling * Work Experience/10	0.019 (0.009)**	0.011 (0.004)***	0.004 (0.008)
Schooling * Cumulative Years Unemployed /10	-0.238 (0.071)***	-0.155 (0.047)***	-0.077 (0.061)
Schooling * Cumulative Years OLF /10	-0.008 (0.053)	-0.049 (0.007)***	-0.066 (0.017)***
Observations	4,228	29,543	4,004
R-squared	0.312	0.350	0.342
Controls:	Cubic Polynomial of Work Experience, Accumulated OLF/Unemployment/Military Years; Uncounted Years; and Years Dummies		

*** p<0.01, ** p<0.05, * p<0.1

White/Huber standard errors clustered at the individual level are reported in parentheses.

Table 1.6: The Effect of Schooling, Experience, and Career Interruptions on Earnings - Leaving School Year as First Year a Responded Left School

NLSY 1979 - Non-Black Males

Dependent Variable: Log Real Hourly Wage

Method: Least Squares

Model	(1)	(2)	(3)
Schooling	0.070 (0.005)***	0.070 (0.005)***	0.057 (0.006)***
Schooling * Work Experience/10	0.018 (0.003)***	0.013 (0.003)***	0.014 (0.003)***
Schooling * Cumulative Years Unemployed /10	-0.151 (0.027)***	-0.137 (0.027)***	-0.167 (0.028)***
Schooling * Cumulative Years OLF /10	-0.003 (0.013)	-0.000 (0.013)	-0.003 (0.014)
Schooling * Tenure Years/10		0.009 (0.001)***	
AFQT * Work Experience/10			0.024 (0.008)***
Observations	38,267	37,665	36,571
R-squared	0.311	0.318	0.325
Tenure	No	Yes	No
AFQT	No	No	Yes

Other controls: Cubic Polynomial of Work Experience, Cumulative Years OLF/Unemployment/Military; Uncounted Years; and Years Dummies

*** p<0.01, ** p<0.05, * p<0.1

White/Huber standard errors clustered at the individual level are reported in parentheses.

Note: Different from the other results, in this table I define year of leaving school as the first year a responded has left school. See section 1.3.1 for details. AFQT is normalized to have a standard deviation of 1. Difference in the number of observations between models (1), (2) and (3) is due to 602 observations of individuals with missing tenure and 1,696 observations of individuals with missing AFQT information.

Table 1.7: The Effect of Unemployment and Schooling on Earnings by Timing of Unemployment

NLSY 1979 - Non-Black Males, 1983-2010
 Dependent Variable: Log Real Hourly Wage
 Method: Least Squares

Model	(1)	(2)	(3)
Schooling	0.104 (0.004)***	0.080 (0.004)***	0.091 (0.005)***
Schooling * Work Experience/10		0.021 (0.003)***	0.016 (0.004)***
Weeks spent unemployed/52			
Last year	-0.219 (0.026)***	-0.229 (0.025)***	0.448 (0.142)***
2 years ago	-0.13 (0.023)***	-0.14 (0.023)***	0.011 -0.133
3 years ago	-0.085 (0.021)***	-0.097 (0.021)***	0.163 -0.129
4 years ago	-0.101 (0.021)***	-0.112 (0.021)***	0.121 -0.133
5 years ago	-0.112 (0.022)***	-0.124 (0.022)***	0.219 (0.128)*
Schooling * Weeks spent unemployed/52			
Last year			-0.056 (0.012)***
2 years ago			-0.012 (0.011)
3 years ago			-0.022 (0.010)**
4 years ago			-0.020 (0.011)*
5 years ago			-0.029 (0.011)***
Observations	31,711	31,711	31,711
R-squared	0.303	0.306	0.308
Controls: Cubic Polynomial of Work Experience, Weeks spent OLF in each of the past 5 year (and their interaction with schooling in model 3), Uncounted Years and Years Dummies			
*** p<0.01, ** p<0.05, * p<0.1			
White/Huber standard errors clustered at the individual level are reported in parentheses.			

Note: The sample is restricted to observations 5 years after an individual's leaving school. Weeks spent in each labor status are constructed using annual aggregation of the week-by-week records.

Table 1.8: The Effect of Schooling, Experience, and Career Interruptions on Earnings, Individual Fixed Effect

NLSY 1979 - Non-Black Males

Dependent Variable: Log Real Hourly Wage

Method: Fixed Effects

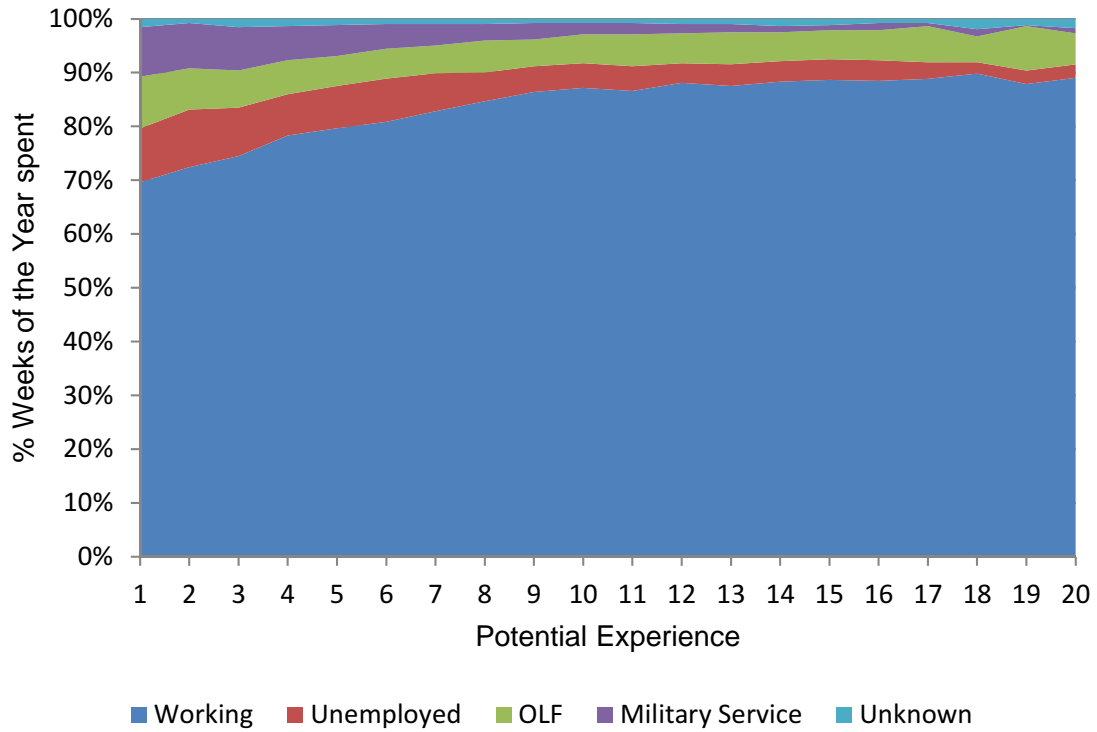
Model	(1)	(2)	(3)
Schooling * Work Experience/10	0.020 (0.003)***	0.015 (0.004)***	0.013 (0.004)***
Schooling * Accumulated Unemployment Years/10	-0.095 (0.036)***	-0.084 (0.036)**	-0.114 (0.036)***
Schooling * Accumulated OLF Years/10	-0.019 (0.034)	-0.019 (0.033)	-0.034 (0.035)
Schooling * Tenure Years/10		0.007 (0.005)	
AFQT * Work Experience/10			0.025 (0.007)***
Observations	33,707	33,181	31,672
R-squared	0.206	0.210	0.215
Tenure	No	Yes	No
Other controls:	Cubic Polynomial of Work Experience, Accumulated OLF/Unemployment/Military Years; Uncounted Years; and Years Dummies		

*** p<0.01, ** p<0.05, * p<0.1

White/Huber standard errors clustered at the individual level are reported in parentheses.

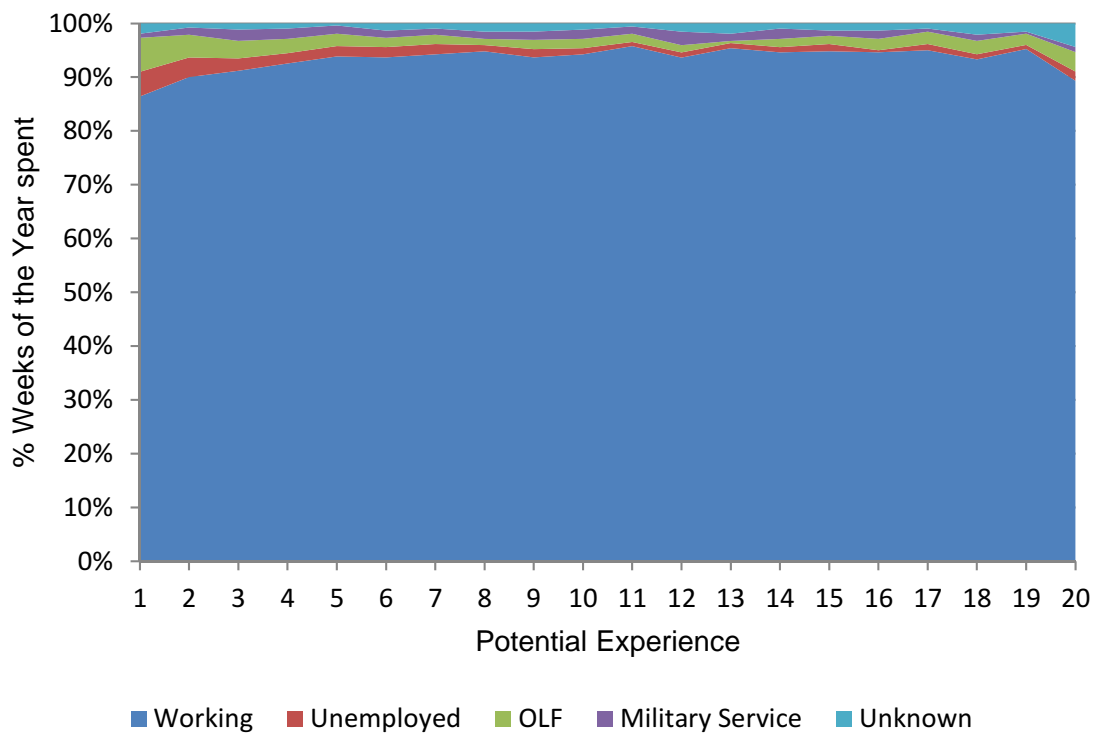
Note: AFQT is normalized to have a standard deviation of 1. Difference in the number of observations between models (1) , (2) and (3) is due to 526 observations of individuals with missing tenure and 1,545 observations of individuals with missing AFQT information.

Figure 1.1: Employment Attachment over the Life-Cycle - High School Graduates



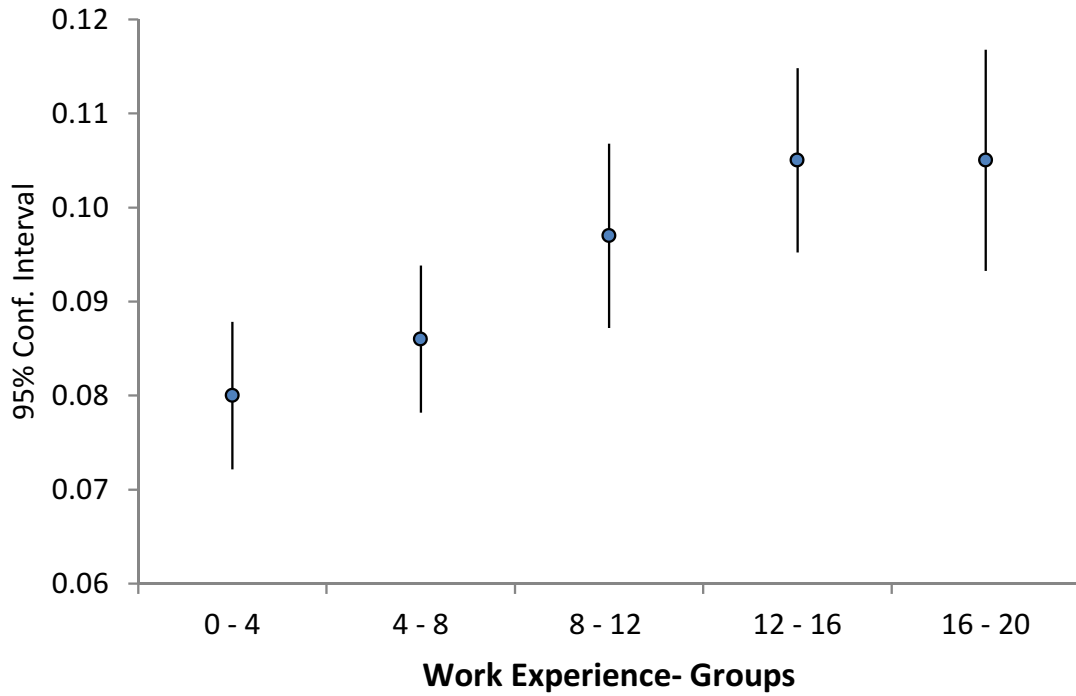
Note: Sample is restricted to observations after an individual left school. Weeks spent in each labor status are constructed using year aggregation of the week-by-week records.

Figure 1.2: Employment Attachment over the Life-Cycle - BA or More Graduates



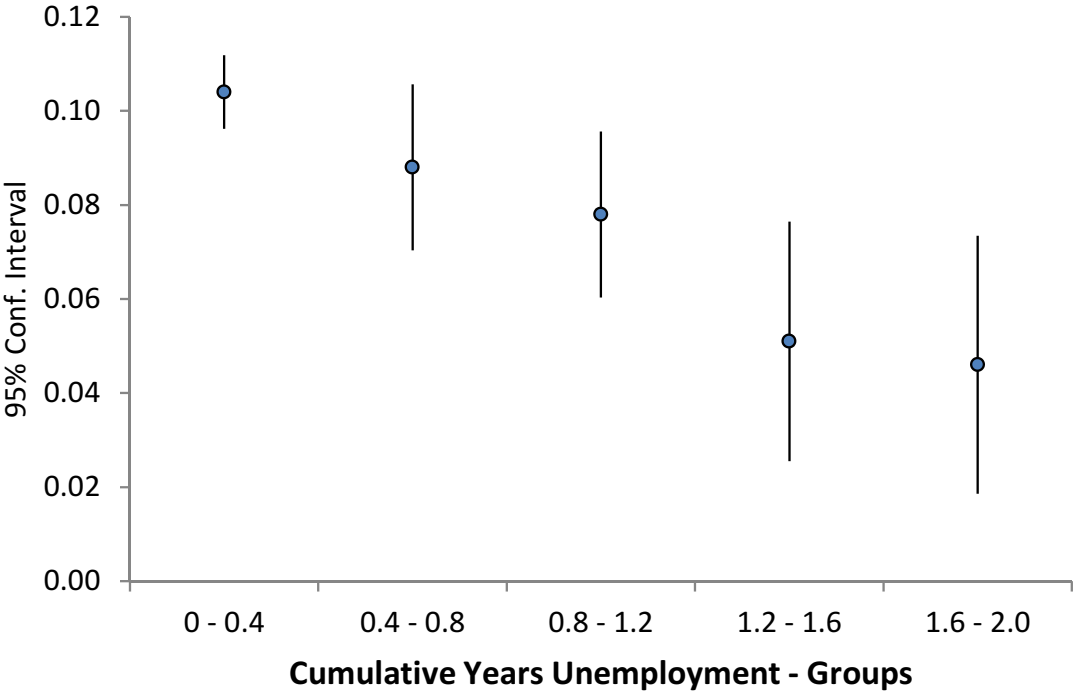
Note: Sample is restricted to observations after an individual left school. Weeks spent in each labor status are constructed using year aggregation of the week-by-week records.

Figure 1.3: Earnings Coefficient on Schooling by Work Experience Groups



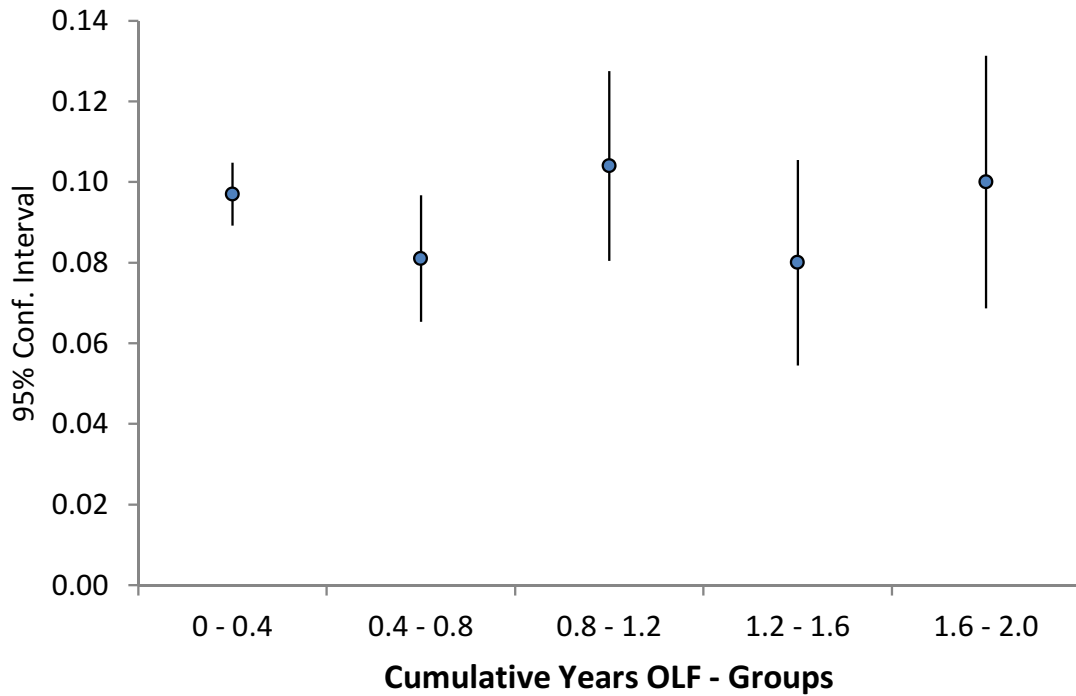
Note: Each circle represents the effect of schooling estimated by linear least squares within each of the 5 Work Experience groups. The controls used in the regressions are the same as those presented in column (2) of table 1.4. Confidence intervals are calculated using White/Huber heteroscedasticity standard errors clustered at the individual level.

Figure 1.4: Earnings Coefficient on Schooling by Cumulative Years Unemployment Groups



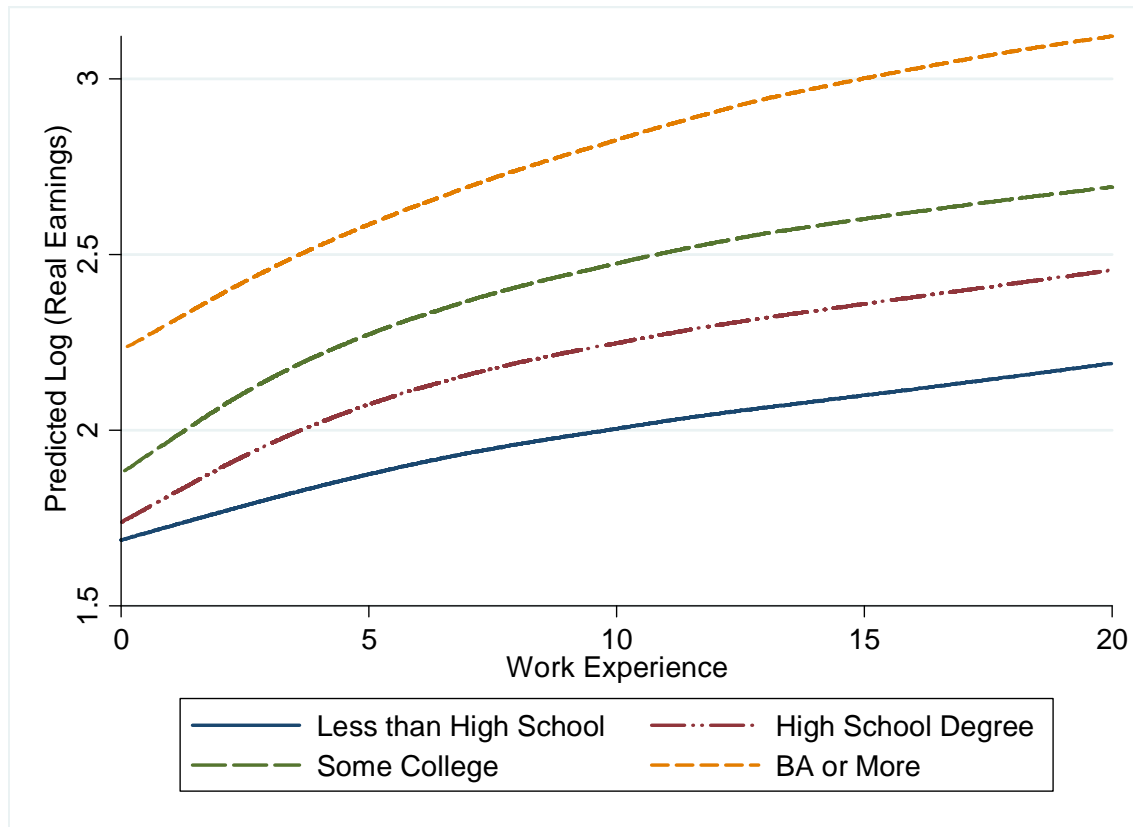
Note: Each circle represents the effect of schooling estimated by linear least squares within each of the 5 Cumulative Years Unemployed groups. The controls used in the regressions are the same as those presented in column (2) of table 1.4. Confidence intervals are calculated using White/Huber heteroscedasticity standard errors clustered at the individual level.

Figure 1.5: Earnings Coefficient on Schooling by Cumulative Years OLF Groups



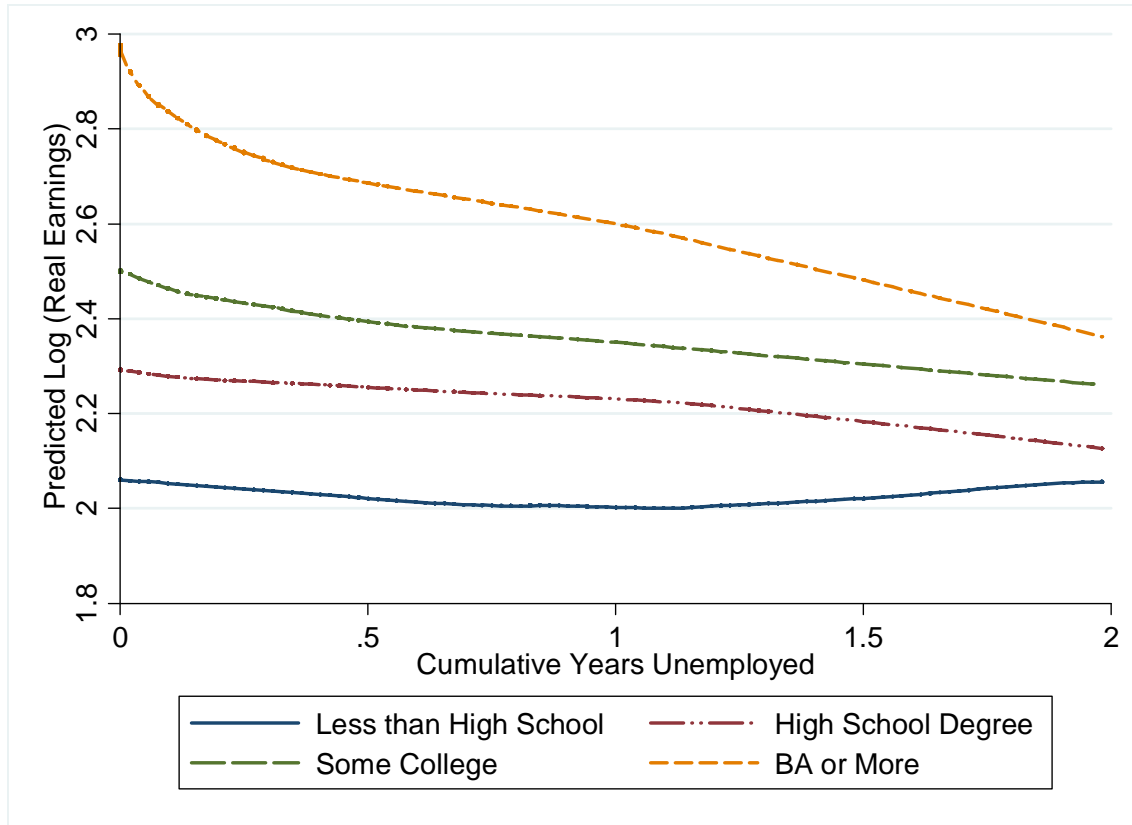
Note: Each circle represents the effect of schooling estimated by linear least squares within each of the 5 Cumulative Years OLF groups. The controls used in the regressions are the same as those presented in column (2) of table 1.4. Confidence intervals are calculated using White/Huber heteroscedasticity standard errors clustered at the individual level.

Figure 1.6: Log Earnings - Work Experience Profile



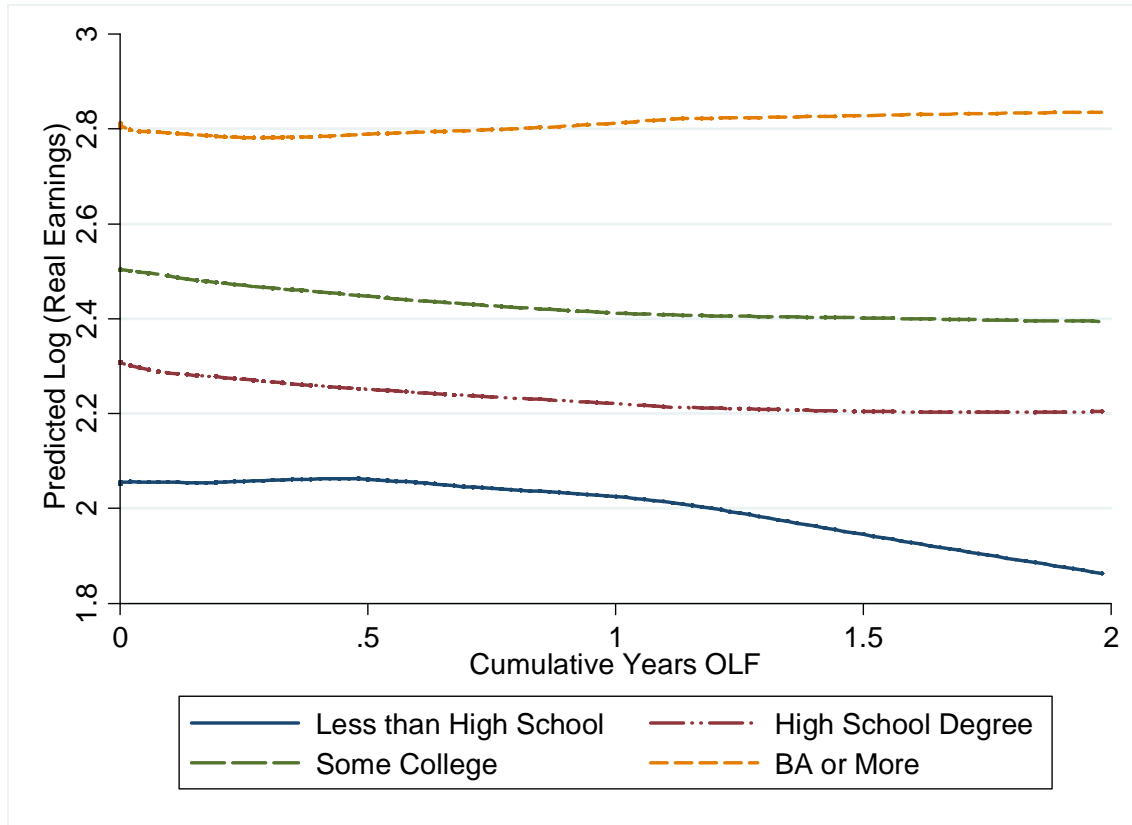
Note: The lines plot the predicted values from a locally weighted regression of log hourly earnings on work experience using a 0.5 bandwidth by each educational group. See section 1.3.2 for details.

Figure 1.7: Log Earnings - Cumulative Years Unemployed Profile



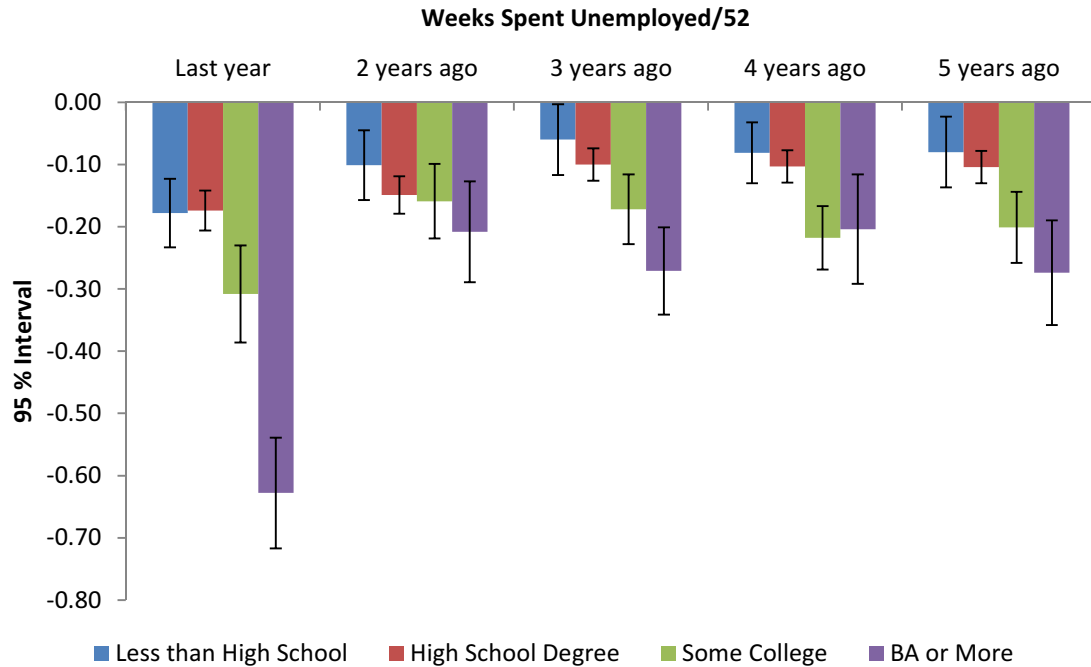
Note: The lines plot the predicted values from a locally weighted regression of log hourly earnings on cumulative years unemployed using a 0.25 bandwidth by each educational group. See section 1.3.2 for details.

Figure 1.8: Log Earnings - Cumulative Years OLF Profile



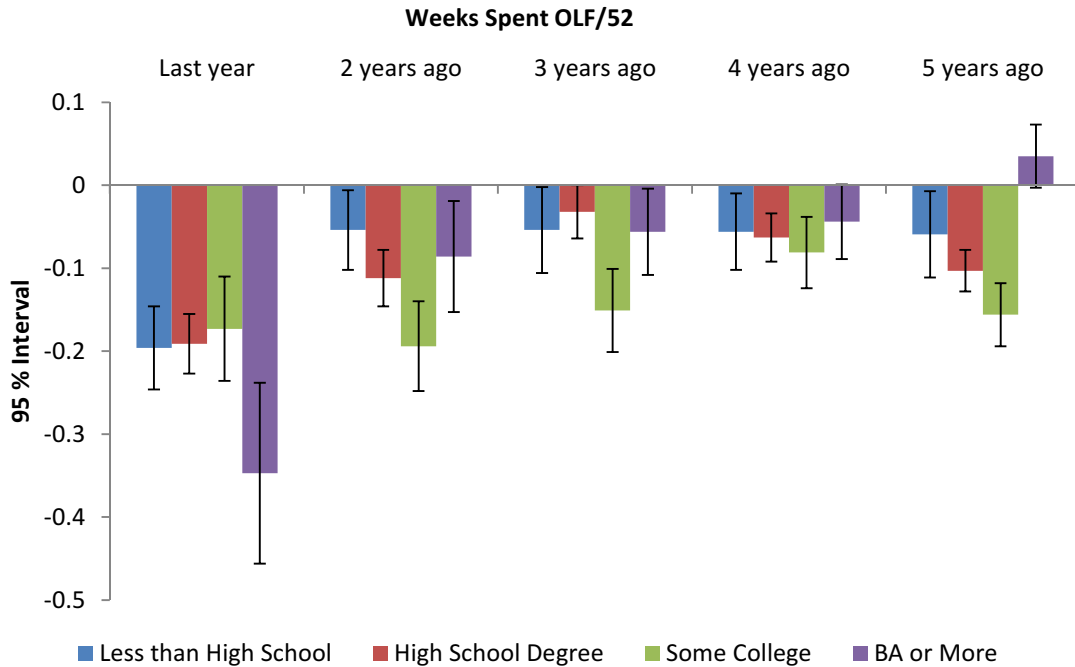
Note: The lines plot the predicted values from a locally weighted regression of log hourly earnings on cumulative year OLF using a 0.25 bandwidth by each educational group. See section 1.3.2 for details.

Figure 1.9: The Effect of Unemployment on Earnings by Timing of Unemployment



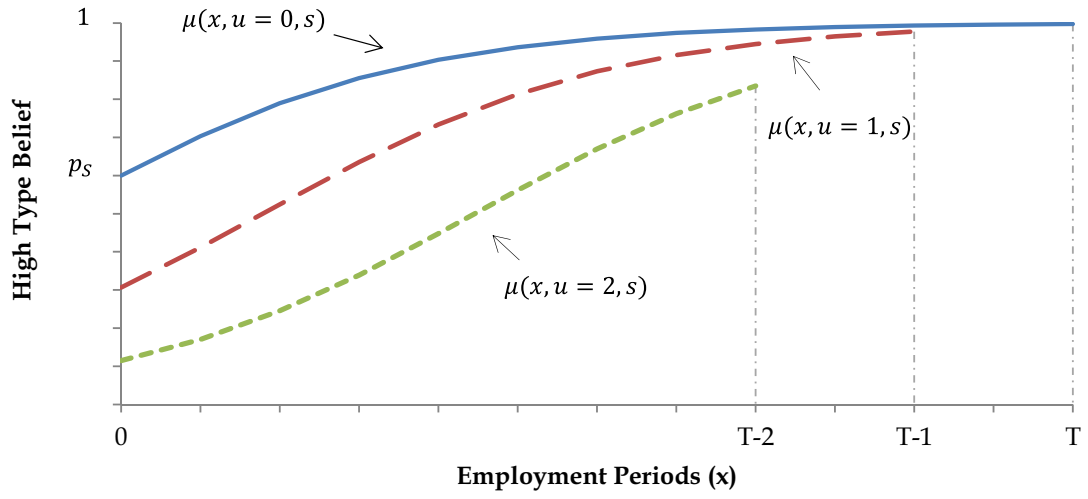
Note: Each bar represents the effect of weeks unemployed in each of the past 5 years conditional on weeks unemployed in the other 4 years. The model is estimated by linear least squares. The controls used are OLF periods, cubic polynomial of work experience, cumulative years military service; uncounted years, and years dummies. Confidence intervals are calculated using White/Huber heteroscedasticity standard errors cluster at the individual level. See section 1.3.2 for details.

Figure 1.10: The Effect of OLF on Earnings by Timing of OLF periods



Note: Each bar represents the effect of weeks OLF in each of the past 5 years conditional on weeks OLF in the other 4 years. The model is estimated by linear least squares. The controls used are unemployment periods, cubic polynomial of work experience, cumulative years military service; uncounted years; and years dummies. Confidence intervals are calculated using White/Huber heteroscedasticity standard errors cluster at the individual level. See section 1.3.2 for details.

Figure 1.11: High Type Belief



CHAPTER II

Employer Learning, Statistical Discrimination and University Prestige⁴⁰

2.1 Introduction

Labor markets are characterized by incomplete information on workers' productivity (Spence, 1973). There are some characteristics of workers, such as labor market ability, that are important for performance on the job but are not easily observable by employers. In this context, firms often have to make judgments on workers' unobservable quality on the basis of the available information. Within this framework, statistical discrimination is defined as employers using group identity of workers to infer their unobservable quality.

The most traditional group identity studied in the statistical discrimination context is race (Phelps (1972) and Aigner and Cain (1977)). In this literature, the racial wage gap is justified not because employers are prejudiced against a particular race but because they use race identity to predict the unobservable quality of workers. More recently, evidence was found that firms use schooling (Farber and Gibbons (1996), Altonji and Pierret (2001) and Lange

⁴⁰This chapter was written with Paola Bordon.

(2007)) or information on lay-offs (Gibbons and Katz (1991) and Hu and Taber (2011)) to statistically discriminate workers.

In this paper we study a new dimension of statistical discrimination: we investigate if firms use the prestige of the university attended by a worker to predict his or her unobservable labor market quality. We believe that college prestige satisfies the typical features of group identity that might be used for statistical discrimination for two main reasons. First, this information is easily accessible to firms: workers use the university name in their resumes and prestigious universities are widely recognized in the labor market. Second, there is evidence that more talented individuals attend more prestigious universities (Hoxby (1998) and Dale and Krueger (2002)). Overall, elite universities have a very competitive application process and tend to select higher quality candidates.⁴¹ Within this framework, it is natural to believe that firms use university prestige in order to infer the unobservable labor market quality of workers.

In order to test if employers use university prestige as a signal of workers' unobservable quality, we rely on the statistical discrimination and employer learning (EL-SD) literature (Altonji and Pierret (2001)).⁴² The underlying assumption is that the imperfect information about a worker's quality tends to disappear with time. At the early stages, firms assess workers on the basis of easily observable variables that are correlated with their unobservable quality. As a worker gains experience in the labor market, employers weigh these characteristic with other information that becomes available, such as references and on-the-job performance. If employers use a characteristic to statistically discriminate a worker in the

⁴¹As it will become clear later, the underlying assumption is that universities are better at screening candidates than firms.

⁴²Other important papers in this literature include Lange (2007), Schönberg (2007), Arcidiacono et al. (2010), and Mansour (2012).

early stage of his career, this information should become less important for earnings as a worker reveals his true productivity with time.

This paper uses data from *Futuro Laboral* of the Chilean Ministry of Education. This data satisfies the purpose of the paper for several reasons: first, it follows different cohorts of college graduate workers from Chile in their first years in the labor market, the period in which most of the employer learning happens (Lange (2007)). Second, the data presents information on labor market outcomes such as earnings from administrative data and we can identify workers that graduate from elite universities. Finally, the data contains information on the scores of the centralized admission test to universities in Chile. As it will be clear later, this information will be used to construct the running variable in the regression discontinuity test we suggest.⁴³

We take advantage of the centralized admission process to college in Chile to propose a statistical discrimination test based on regression discontinuity design (RD). Using information on the admission test scores we are able to identify workers who were just above or just below the admission thresholds to the two most prestigious universities in Chile. We suggest an EL-SD test that compares the earnings' dynamics between these two group of workers as they gain experience in the labor market. The test predicts that if firms use university prestige to statistic discriminate workers: i) individuals barely admitted to the most selective universities in Chile should be paid substantially more than those barely rejected when they graduate from college; ii) the wage differential between these two group of workers should

⁴³Kaufmann et al. (2012) and Hastings et al. (2013) are two recent papers that have also explored the regression discontinuities generated by the centralized admission process to universities in Chile. Kaufmann et al. (2012) looks at effect of graduating from a elite university on marriage outcomes and Hastings et al. (2013) studies labor market returns to college admission. None of these papers explore how the selective university wage premium changes throughout a worker's career, that is the main object of interest of this paper.

shrink as individuals progress in their career.

The idea for the test is similar to the one presented by Altonji and Pierret (2001). Employers do not observe admission test scores but they know that prestigious universities admit on average better candidates. If employers use the selectivity of a university as a signal of worker's inherent ability, individuals just above the admission cutoff must be better paid than those just below the admission cutoff. Nevertheless, the EL-SD model proposes that employers learn a worker's unobservable quality with time. In consequence, the signal associated with graduating from a prestigious university should become less important for earnings and the wage differential between workers similar pre-college characteristics should shrink with time. We find evidence for statistical discrimination on the basis of university prestige. We estimate that workers just above the admission cutoff to the two most prestigious universities in Chile earn on average 12% more than those just below the cutoff in the first year after graduation. However, this wage premium tend to decrease by 2 percentage points by year of experience in the labor market, to the point that we cannot reject a zero earnings differential between these two groups of workers 4 years after their graduation. We also take into consideration the fuzziness of the the regression discontinuity design, meaning that not all students admitted to a prestigious university in Chile attend such university, to estimate the local average effect of graduating from a prestigious university on earnings. We estimate a 19% wage premium for recent graduates of the two most prestigious university in Chile. However, this wage premium decreases by 3 percentage points per year of experience.

Based on these findings, this paper contributes to different dimensions of the existing literature. First, this paper is a contribution to the EL-SD literature because we study statistical discrimination on the basis of a different group identity. While there is an extensive literature that analyzes the use of race, gender, and schooling, we are one of the first papers to study

whether firms use prestige of college to statistically discriminate workers.⁴⁴ Furthermore, to the best of our knowledge, our paper is the first to propose an employer learning-statistical discrimination test based on a regression discontinuity design. We present issues with the traditional test proposed by Altonji and Pierret (2001) if employers statistically discriminate workers on the basis of characteristics that are not present in the data and are correlated with graduating from a prestigious university, such as family social-economic background. We also demonstrate that the regression discontinuity test we propose is robust to such bias. Second, we contribute to the literature which studies the effect of graduating from an elite university on labor market outcomes. There is an extensive series of papers that estimate the returns to graduating from a selective university on earnings (Brewer et al. (1999), Hoxby (1998), Dale and Krueger (2002), and Black and Smith (2006)), including papers that have used a regression discontinuity design (Saavedra (2008) and Hoekstra (2009)). The overall finding is that there is a positive effect of graduating from a prestigious college on earnings.⁴⁵ While there is big effort in the literature to overcome the selection bias associated to attending a prestigious university, little attention has been given to the mechanisms that generate the college selectivity wage premium.⁴⁶

Different from past work, in this paper we shed some light on the reasons for why workers

⁴⁴To the best of our knowledge, Lang and Sniver (2011) and Hershbein (2013) are the two other papers that have addressed this issue. Lang and Sniver have a similar approach to estimate how returns to attending a elite university in Israel changes with labor market experience. However, the authors are unable to properly exploit the regression discontinuity in the college admission.

⁴⁵The only exception is Dale and Krueger (2002) who find no wage premium from attending a selective college. It is interesting to note that the authors estimates the wage premium approximately 15-19 years after a worker's graduation from college. The zero effects for individuals with similar pre-college characteristics later in advanced age does not contradict the empirical findings of this paper.

⁴⁶Pop-Eleches and Urquiola (2013) is one of the few papers that have discussed the benefits from attending a higher quality school. Their paper address the behavior effect of students, parents and teachers in response to a student admission to better secondary school. Nevertheless, the paper has little to say about the impact of attending a higher quality school on labor market outcomes.

from prestigious universities receive higher wages after graduation. On one hand, attending a selective university could be associated with receiving better instruction and having more accomplished peers. In this context, prestigious universities have an advantage of increasing a worker's productivity in comparison to less prestigious universities. On the other hand, the main effect of attending a selective university might be to signal to employers an unobservable inherent ability of a worker. In this context, the extra value added from a selective college education might not be significantly higher than that from a less prestigious university. Using the regression discontinuity test we propose we are able to disentangle the signaling effects from the value added effect of graduating from a prestigious university. Our finding of a rapid decrease in the elite college premium for workers with similar pre-university characteristics is evidence that signaling mechanisms are stronger than productivity mechanisms. In particular, the fact that we cannot reject a wage differential between workers just above and below the admission cutoff after 4 years in the labor market suggests that the value added from the two most prestigious university in Chile is not significantly different from the less prestigious schools.

2.2 Institutional Framework

Higher education in Chile comprises three types of institutions: Universities, Professional Institutes (IPs), and Technical Formation Centers (CFTs). Universities provide the highest degree of learning, combining teaching, research and outreach activities; they teach accredited degree programs (2.5 to 4 years) and award academic degrees (5 to 7 years). Professional Institutes are in charge of granting professional degrees other than those awarded by universities, and they are also authorized to grant higher education technical degrees in areas where

this is required. Technical Formation Centers are intended to equip higher level technicians with the competencies and skills needed to respond to the needs of industry in the public and private sectors.

Universities can be divided into two main categories: traditional and non-traditional institutions. Traditional institutions comprise the oldest and most prestigious universities created before 1981, and those institutions that derived from the old universities (created after 1980). Traditional establishments consist of 25 fully autonomous universities coordinated by the Council of Chancellors of Chilean Universities (CRUCH) and are eligible to obtain partial funding from the state. They employ a single admission process: the University Selection Test (PAA)⁴⁷. This test is made up of three compulsory sub-tests including language, mathematics, and history and geography of Chile. Additionally, depending on which programs they are planning to apply to, students may be required to take the following specific PAA tests: advanced mathematics, physics, chemistry, biology, and history.

The time-line of the admission process into traditional universities happens as described in figure 2.12. First, students take the PAA test and after receiving their score they make their application choices. Students apply to a major and university (or program) simultaneously and can only apply to 8 programs, ranking them up by preferences. The only criterion used for admission in the traditional universities is the score in the PAA. This final admission scores consists of a weighted average of the compulsory and major specific tests and high school GPA, with each program setting its specific PAA weights.⁴⁸ The number of vacancies for each program is announced before the application process and programs fill their vacancies

⁴⁷In 2004 the university selection test was modified and it is now called PSU.

⁴⁸For example, engineering in a prestigious university requires 20% of mathematics, 10% of language, 10% of history, 20% high school GPA, 30% specifics mathematics, and 10% physics. The final score to the same major in a different university might requires different weights.

solely based on the final weighted scores. The admission score cutoff is defined by the score of the last student admitted into a program and it is not known before the application decisions and therefore students cannot manipulate which side of the cutoff on they fall on.⁴⁹

Non-traditional universities were created after 1981, have no state financial support and might not necessarily use the PAA score to select their incoming students. Nevertheless, the anecdotal evidence is that the majority of students willing to attend higher education in Chile take the PAA at the end of high school independent of the university they are planning to attend. The test is relatively inexpensive and administrated throughout the country.

All higher education institutions charge tuition and fees. However, for those students enrolled in one of the traditional universities, solidarity credits and scholarships are available. Competition in these markets, particularly for undergraduates, is often geographically circumscribed to local and regional markets, and it can be more or less intense depending on the institution. As of 2001, the Chilean higher education system consisted of 60 universities (25 traditional universities and 35 new private universities without direct public subsidy), 42 professional institutes (all of them private), and 117 private technical formation centers.

The increasing enrollment in higher education has led to an increasing number of graduates in the last two decades. In 1995, 24,400 graduates entered the labor market, whereas in 2000 around 42,000 graduates did, and in 2005, 71,170 new graduates were entering the job market. This means than in ten years the number of graduates has almost tripled. Traditional universities have more than doubled the number of graduates they produce, but private universities have increased by 6.7 times their number of graduates.

⁴⁹Students could use the admission score cutoff of previous years as a reference. Given the variation of the admission cutoff overtime and the possibility to apply to 8 different programs, we believe that students with marginal scores to be admitted in prestigious university tend to to apply to these competitive programs.

2.3 Data

The data to be used in the study comes from *Futuro Laboral*, a project of the Ministry of Education of Chile that follows individuals over the first years of labor market experience after graduating from higher education programs. The panel data set matches tax returns with transcripts of students' majors and the institutions they graduated from. The unit of analysis concerns only those who graduate from both traditional and non-traditional universities; those who have stopped studying or did not continue their studies after graduating from high school are not in the sample. Income information is available between the years 1996 and 2005. We have data for the 1995, 1998, 2000 and 2001 graduating classes.⁵⁰

The information provided by the Internal Revenue Service (SII) comprises age, sex, name of the institution that individuals graduated from, major, the year of graduation, annual income reported in tax returns, city or cities of employment, number of employers and economic sector. The raw data contains every worker in Chile that had positive earnings between 1996 and 2005, even those who exempt from tax.⁵¹ ⁵² For a random sub-sample, the Ministry of Education gathers information about the PAA score, high school grades and the institutions students graduated from high school. As the PAA scores have an important role in both the traditional EL-SD and regression discontinuity analysis, we restrict our study to this sub-sample.

⁵⁰Note that the cohorts are observed for different length of time. For example, while we observe 10 years of labor market experience for the 1995 graduation class, we only 4 years of labor market experience for the 2001. Unfortunately, the project was deactivated and the income data for more recent years was not collected.

⁵¹Note that in Chile, married couples must fill their taxes separately.

⁵²A concern is that part of the individual from prestigious universities might go to graduate school after finishing their baccalaureate studies and therefore would be omitted in the earnings sample. However, the fraction of workers that go to graduate school in Chile is very low. Using data from the National Socioeconomic Characterization Survey in the year 2000, we find that only 0.65% of 25-34 years old individuals with a bachelor degree were enrolled in graduate school or had obtained a graduate degree.

The wage measured in the sample is the annual income that comes from jobs and services provided by the individual.⁵³ We use consumer price index (IPC) as a deflator to compute real wages. The experience variable is computed as the number of years an individual has income and has paid taxes after graduation. The final sample consists of 58,477 individuals and 322,688 observations.

We divide universities into two groups: selective and non-selective universities. The selective universities comprises two of the oldest and most prestigious universities in the country. These schools attract students with the highest PAA scores and therefore are the most selective schools in the country. The programs of these two universities have also been consistently ranked among the highest in Chile and their prestige is well recognized nationwide.⁵⁴ See Table 2.9 for descriptive statistics regarding these two groups. As expected, selective universities have on average higher scores in Math and Language components of the PAA tests, and their students have higher high school grades. We also see that 11% of selective universities students went to a private high school, compared to 7% from non-selective universities. We also plot in the distribution of language and math PAA scores for college graduates from selective and non-selective universities on figures 2.13 and 2.14 respectively. One can see from the figures that the language and math scores of graduates from selective universities are concentrated at the higher end of the distribution. Finally, we show in Table 2.10 that workers from the two selective universities have on average higher earnings than those from the less prestigious schools.

⁵³We do not have information on weeks or hours worked in the sample and for this reason we cannot explore how much of the annual income increase is due to changes in hours or week of work. Nevertheless, workers with a bachelor degree in Chile present both a high employment attachment and the majority work full time. Using the National Socioeconomic Characterization Survey in the year 2000, we find that 86.7% of 25-34 years old individuals with a bachelor degree work are employed in the period of the interview and from those, 88% work more than 35 hours per week.

⁵⁴Due to a confidentiality agreement with the Ministry of Education, we cannot provide the name of these two institutions.

2.4 Regression Discontinuity Test

In order to provide evidence for statistical discrimination based on college prestige, we use a regression discontinuity (RD) design. The basic idea is to compare how the earnings of those just above and just below the cutoff for admission to the most selective universities in Chile change as workers accumulate experience in the labor market. The identification assumption is that other factors that could affect earnings are continuous at the admission cutoff and students have limited power to manipulate on which side of the admission cutoffs they might fall.⁵⁵ Furthermore, we assume that employer do not have access to the test scores that a prestigious university uses in their admission process.

2.4.1 Employer Learning Statistical Discrimination Model

The standard employer learning model specifies the log-productivity of a college graduate worker i with experience level t :

$$y_{it} = rs_i + \alpha_1 q_i + \lambda z_i + \eta_i + H(t) \quad (2.17)$$

where s_i captures information that is available to both employers and researchers. In this paper, s_i is defined as an indicator if a worker graduated from a prestigious university or not. The variable q_i describes information available to employers and not present in the data, such as family social economic background, z_i is a characteristic present in the data but not available to employers and η_i is a measure of a worker's inherent ability that is not available

⁵⁵Students can retake the test next year, but they cannot retake the test after they got their results the same year, which decreases the probability of manipulation.

in the data or to employers. Finally, $H(t)$ describes the relation between log-productivity and experience and does not depend on the other variables of the model.

In the absence of information on z_i and q_i , employers form expectations based on other observed characteristics of workers. Altonji and Pierret (2001) assume that these conditional expectations are linear on s and q :⁵⁶

$$z = \mathbb{E}[z|s, q] + v = \gamma_1 q_i + \gamma_2 s + v$$

$$\eta = \mathbb{E}[\eta|s, q] + e = \alpha_2 s + e$$

where v and e are scalar with mean zero and uncorrelated with s and q by construction. Under this assumption, one can characterize the expected value of y given information on s and q :

$$\mathbb{E}[y|s, q] = (r + \lambda\gamma_2 + \alpha_2)s + (\alpha_1 + \lambda\gamma_1)q + H(t)$$

In the traditional EL- SD model, employers have access to a noisy measure of a worker's productivity after each period that an individual spend in the labor market:

$$\tilde{y}_{i\tau} = y_{it} + \varepsilon_{i\tau}$$

where the noise ε_t is independent of all the variables of the model. As in Altonji and Pierret (2001), employers share equal information about workers, labor markets are competitive and there is a spot market for labor services. As a consequence, wages are equal to the expected productivity of a worker, given the information available to employers at each period.

⁵⁶A normalization allows suppressing q in the second expectation.

$$W_{it} = \mathbb{E}[\exp(y_{it})|s, q, \tilde{y}_{i0}, \dots, \tilde{y}_{it-1}]$$

Lange (2007) assumes that ε_t is independently, identically and normally distributed with a finite variance. Under this assumption, the process of updating the expectations of employers have a very simple structure an the log-wage process can be represent by:

$$w_{it} = (1 - \theta_t)\mathbb{E}[y|s, q] + \theta_t \frac{1}{t} \sum_{\tau=0}^{t-1} \tilde{y}_\tau + \tilde{H}(t) \quad (2.18)$$

where $\tilde{H}(t)$ is a linear transformation of $H(t)$ and θ_t is a function of the variances of and $\varepsilon_{i\tau}$, s and q . Furthermore $\theta_0 = 0$ and θ_t strictly increases with t converging to 1 as t goes to infinite.⁵⁷ This expression demonstrates that as a worker progress in his career, employer weight less their initial believe on a worker's productivity based on s and q , and weight more the new information that becomes available during a worker's career.

Traditional EL-SD Test

The object of interest in the traditional employer learning model is the linear projection of the log-wage w_{it} on s , z and t .

$$\mathbb{E}^*[w_{it}|s, z, x] = b_{sx}s + b_{zx}z + \tilde{H}(t)$$

Without lost of generality, one can define the the projections of the unobservable variables (q, η) on the observable variables (s, z) :

$$q = \gamma_3s + \gamma_4z + u_1$$

⁵⁷See Lange (2007) for the formal derivations of these parameters.

$$\eta = \gamma_5 s + \gamma_6 z + u_2$$

Using the independence of ε_{it} to all the variables of the model, Lange (2007) show that the coefficients of the projections:

$$b_{st} = (1 - \theta_t)b_{s0} + \theta_t b_{s\infty} \quad (2.19)$$

$$b_{zt} = (1 - \theta_t)b_{z0} + \theta_t b_{z\infty} \quad (2.20)$$

where, as discuss before, $\theta_0 = 0$ and $\lim_{t \rightarrow \infty} \theta_t = 1$. The traditional EL-SD test consists in estimating how b_{st} and b_{zt} change with experience level t . Indeed, Altonji and Pierret (2001) propose that if firms statistically discriminate workers on the basis of s and if z is positively related to s , one should observe that b_{st} falls with t and b_{zt} should rise with t .

Furthermore, under the assumptions above Lange (2007) shows that:

$$b_{s0} = \underbrace{r}_A + \underbrace{\alpha_1 \gamma_3}_B + \underbrace{\alpha_2 + \lambda(\gamma_2 + \gamma_1 \gamma_3)}_C \quad (2.21)$$

$$b_{z0} = \underbrace{(\alpha_1 + \lambda \gamma_1) \gamma_4}_D \quad (2.22)$$

where the coefficient b_{s0} represents the relation between graduating from a prestigious university and wages in the beginning of a workers career. The first term A captures the direct effect of attending a prestigious university on productivity. The second term B represents the direct impact of q on wages and the fact that q is not present in the data but it is correlated to s . This can be interpreted as the traditional omitted variable problem associated

with estimating the returns to graduating from a prestigious university (Dale and Krueger (2002)). It captures the relation between any variable that affects wages, is correlated to graduating from a prestigious university and it is not present in the data. Finally, the term C reflects the fact that employers do not observe η and z in the beginning of a worker's career, but are aware of their relation with s . Therefore, employer use s as a signal of unobservable components of a worker's productivity. In the same way, the relation between z and the log wages of a worker in the beginning of his career is given by the coefficient b_{z0} . As employers do not observe z , this coefficient only captures the fact that we are omitting q from the linear prediction and that z and q are correlated.

$$b_{s\infty} = \underbrace{r}_E + \underbrace{\alpha_1\gamma_3 + \gamma_5}_F \quad (2.23)$$

$$b_{z\infty} = \underbrace{\lambda}_G + \underbrace{\alpha_1\gamma_4 + \gamma_6}_H \quad (2.24)$$

The coefficients $b_{s\infty}$ and $b_{z\infty}$ represent the relation between s and z respectively with wages as $t \rightarrow \infty$ and $\theta_t \rightarrow 1$. As before, E represents the direct effect of graduating from a prestigious university on wages. The coefficient F captures the fact that η and q have an impact on long-run wages, are related to s but they are omitted in the linear prediction because they are not observed in the data. Note that F is different from the term B because firms only learn η with time. In the same way, the term G captures the direct impact of z on productivity and H captures the correlation of z to the omitted variables η and q .

One important issue that has been omitted from the employer learning literature (Altonji and Pierret (2001) and Lange (2007)) is how the correlation between s and the unobservable

factor q can affect the conclusions of statistical discrimination test. This issue arises if firms statistically discriminate workers on the basis of variables that are not observed in the data, such as family social economic background, that are correlated to graduating from prestigious university. In this situation, the traditional employer learning test might suggest that employers statistically discriminate a worker on the basis of university prestige, when in fact firms might be using family social economic status as a signal of a worker's unobservable characteristics.

In order to give some perspective of the issue, we analyze the extreme case where s is not correlated to η and z ($\alpha_2 = 0$, $\gamma_2 = 0$ and $\gamma_5 = 0$). In this situation, employers should not use s as a signal of a worker's unobservable characteristics, and therefore, workers are not statically discriminated on the basis of university prestige. Furthermore, assuming that q is correlated with η and z ($\gamma_4 \neq 0$), and therefore q is used by employer to statistically discriminate workers. Under this assumption, the traditional employer learning test would suggest that firms statistically discriminate workers on the basis of university prestige because $b_{s\infty} < b_{s0}$ and $b_{z\infty} > b_{z0}$. Note, however that this conclusion is being driven by the correlation of s and q , and the fact that employers use q to predict z , which is capture by the term $\lambda\gamma_1\gamma_3$ in equation (2.22).

Regression Discontinuity EL-SD Test

The object of interest of the EL-SD test we propose is how the difference between average log-wages of individuals just above and just below the admission cutoff to a prestigious university changes with experience. Precisely, we define $Dist.Cutoff_i$ as the distance between a student's test score and the admission threshold of a prestigious university. For simplicity, we assume that all students admitted to a prestigious university graduate from this university,

such that $s_i = 1$ if $Dist.Cutoff_i \geq 0$ and $s_i = 0$ otherwise.⁵⁸

The parameter of interest in the paper is:

$$\tau_t = \lim_{Dist.Cutoff \downarrow 0} \mathbb{E}[w_{it}|Dist.Cutoff_i] - \lim_{Dist.Cutoff \uparrow 0} \mathbb{E}[w_{it}|Dist.Cutoff_i] \quad (2.25)$$

that represents local average difference of log-wages by experience levels at the admission cutoff. The employer learning statistical discrimination consists in testing if τ_t decreases with t .

Note that by definition, we have that:

$$\lim_{Dist.Cutoff \downarrow 0} \mathbb{E}[s|Dist.Cutoff_i] = 1 \text{ and } \lim_{Dist.Cutoff \uparrow 0} \mathbb{E}[s|Dist.Cutoff_i] = 0$$

Furthermore, we assume that the distribution of the other variables of the model $\{z_i, q_i, \eta_i\}$ is continuous around the admission cut-offs. In this case, the expected values of these variables just above and just below the admission cutoff are the same:

$$\lim_{Dist.Cutoff \downarrow 0} \mathbb{E}[X|Dist.Cutoff_i] = \lim_{Dist.Cutoff \uparrow 0} \mathbb{E}[X|Dist.Cutoff_i]$$

for $X = q, z, \eta$. Using these two conditions, the assumption that employer do not have access to $Dist.Cutoff_i$, and the the log-wage process derived in (2.18), one can show that:

$$\tau_t = (1 - \theta_t)(r + \lambda\gamma_2 + \alpha_2) + \theta_t r$$

⁵⁸As it will be clear later, this assumptions is not confirmed in the data because some students admitted to a prestigious university decide to attend a less prestigious university (fuzzy regression discontinuity). For simplicity, we ignore this possibility here.

$$= \underbrace{r}_I + (1 - \theta_t) \underbrace{(\alpha_2 + \lambda\gamma_2)}_L \quad (2.26)$$

where θ_t is defined in the same way as above. The regression discontinuity effect of graduating from a prestigious university on wages at experience level x is composed by two terms. The first term I represents the direct effect of s on the workers productivity. The second term L represents the fact that employers do not observe η and z and use s as a signal for these two variables. In other words, if firms statistically discriminate among workers on the basis of university prestige, we have that $L > 0$. However, the signaling term L becomes less important for earnings as firms learn about a workers true productivity, τ_t decreases with t and converges to r as θ_t goes to 1.

There are is an important difference between the regression discontinuity test we propose and the traditional employer learning test: the parameter τ_t does not depend on the relation between s and q . In other words, the regression discontinuity test is robust to the existence of characteristics that could be used for statistic discrimination that are related to graduating from a prestigious university and that are not present in the data. This difference is important because, as discussed above, the traditional EL-SD test might confound statistical discrimination based on family socioeconomic status and statistical discrimination based on college prestige since these factors are intrinsically related and we do not observe family socioeconomic status in the data.

2.4.2 The Admission Process and the RD Design

Our data contains information on the year a student took the PAA test, his or her scores on each component of the test, the college he or she graduated from and the major. We do

not observe application decisions and therefore have to make extra assumptions and sample restrictions to perform the regression discontinuity design. Precisely, we restrict the data to individuals who graduated with engineering, business, medical and law degrees (competitive majors) and assume that these workers would prefer to graduate with these majors in a least prestigious university rather than study a different major in a prestigious university. Under this assumption, we can interpret that workers just above the admission cutoff (competitive major at prestigious universities) are those who were accepted to the highest program of their preference and those below the threshold (competitive major in less prestigious college) are those who were accepted to the second highest program of their preference. We find evidence that this is a plausible assumption. First, these are the programs with highest admission cutoffs and therefore should be top choices of applicants. Second, there is a positive wage differential between workers with the competitive majors in less prestigious university and workers with less competitive major in prestigious university. We interpret this as evidence that students have incentives to study engineering, business, medical or law degree at a less prestigious rather than other major in a prestigious university.

Using additional data on the PAA weights used by these programs in the two prestigious universities we are able to reconstruct the final weighted score for all individuals in the restricted sample.⁵⁹ As a result, we derive $Univ1.Score_i$ and $Univ2.Score_i$ that represents the PAA weighted score of individual i at prestigious university 1 and 2 respectively.

Given the possibility that a student can be accepted in two, one or neither of the prestigious universities, we define the running variable used in the RD as follows:

$$Dist.Cutoff_i = \max\{Univ1.Score_i - Univ1.Cutoff_i, Univ2.Score_i - Univ2.Cutoff_i\}$$

⁵⁹We were only able to obtain PAA weights for years starting in the year 2000. In order to construct final scores for individuals that too the PAA prior to 2000, we assume that programs used the same weights for previous years. The evidence is that programs do not change weights over time.

where $Univ1.Cutoff_i$ and $Univ2.Cutoff_j$ are the admission score cutoffs used by universities 1 and 2 for individual i 's major in the year of application to college. Note that individuals with $Dist.Cutoff_i$ slightly greater than zero were barely admitted to at least one of the two prestigious universities and individuals with slightly lower than zero were barely reject by both schools.⁶⁰

In the RD design we will be interested in the following object:

$$\tau_t = \frac{\lim_{Dist.Cutoff \downarrow 0} \mathbb{E}[w_{it} | Dist.Cutoff_i] - \lim_{Dist.Cutoff \uparrow 0} \mathbb{E}[w_{it} | Dist.Cutoff_i]}{\lim_{Dist.Cutoff \downarrow 0} \mathbb{E}[g_i | Dist.Cutoff_i] - \lim_{Dist.Cutoff \uparrow 0} \mathbb{E}[g_i | Dist.Cutoff_i]}$$

where g_i is an indicator if worker i graduated from an elite university, t measures years of experience in the labor market, and w_{it} is the log(wages) after t years of experience. Note that the parameter τ_t represents the local average treatment effect on earnings after t years of experience for workers around the admission cutoffs that would enroll in a prestigious university if they were admitted (intent-to-treat effect).⁶¹

The employer learning-statistical discrimination RD test we propose consists of estimating if τ_t decreases with t . The test is based on the assumption that the unobserved ability (η_i) is positively correlated to graduating from a selective university but is continuous around the admission cutoff. In this framework, assuming that firms do not observe $Dist.Cutoff_i$, they will use information on college prestige to predict that workers just above the admission cutoff have a higher η_i .⁶² However, the wage differential between those above and below the

⁶⁰Information on program admission cutoffs were collected at the universities websites (late application years) and newspapers (early application years). We find that 4% of individuals in our restricted sample with a prestigious university degree have weighted scores lower than the admission cutoffs. This could be justified by measurement errors in the admission cutoffs and weighs used in the paper or transfers from less prestigious universities. We drop these individuals from the sample used in the RD analysis.

⁶¹For a discussion of the relationship between regression discontinuity design and treatment effects, see Lee and Lemieux (2010)

⁶²Note that in section 4 we also assume that firms cannot observe $Dist.Cutoff_i$. Furthermore, screening workers is expensive and firms learn fast (Lange (2007)), therefore it is not economically attractable to perform ability tests on recent college graduates.

cutoff should decline if firms learn the true distribution of η_i as workers gain experience and therefore should rely less on college prestige to set wages.

2.4.3 Results

We first address the empirical question if the probability of graduating from one of the two prestigious universities in Chile is discontinuous at the admission cutoff. Note that it is possible that individuals with a higher score than the admission cutoffs decided to attend a less prestigious university, which implies that we have a fuzzy regression discontinuity design. Figure 2.15 shows graphically the discontinuity in the probability of graduating from a prestigious university at the cutoff. From the figure, we find that the discontinuity in graduation from a prestigious university is approximately 60 percentage. This means that around 60% of the individuals with PAA scores just sufficiently high for admission choose to attend an elite university. Consequently, being just above the admission cutoff causes a large increase in the probability of graduating from a prestigious university in Chile, which is a necessary condition for the validity of the RD design.

Next, in figure 2.16 we present further evidence for the validity of the RD design. The basic idea is to test if there is a jump at the discontinuity for per-treatment variables that should not be affected by the treatment. Precisely, if being above or below the cut-off is random, we should observe a zero treatment effect on the probability of being female or graduating from a private high school (Imbens and Lemieux (2008)). The figure suggests that there is no discontinuity of these variables around the cutoff. In fact, from a formal test using the same specification in columns (1) to (3) of table 2.11 but using female or private high school indicator as dependent variable, we cannot reject at reasonable levels of significance that

there are zero effects of being above the cut-off on these per-treatment outcomes.⁶³

In order to present evidence of the effects of admission to a selective university on earnings, we plot in figure 2.17 unconditional means of log annual earnings on the vertical axis and the distance from the admission cutoff on the horizontal axis for the first 4 years of labor market experience. The open circles represent 16 points local average and the lines represent linear fits of the data below and above the admission cutoff. The figure shows that there is a jump in earnings in the first year of labor market experience for workers who are just above the cutoff. This discontinuity is consistent with previous literature that finds a significant effect on earnings for being just above the admission cutoff of recent college graduates (Saavedra (2008)). However, as workers gain labor market experience, the discontinuity in earnings tend to decrease to the point that there is no apparent difference in terms of earnings between workers just above and just below the cutoffs four years after graduation. In addition to that, we observe that workers tend to be paid more in accordance with their weighted score as they accumulate experience in the market.

Table 2.11 presents further statistical evidence for discontinuity in earnings at the admission cutoff. In columns (1) to (3) of panel A of the table, we show that workers above the admission cutoff have on average 6-8% higher earnings than just below the admission cutoff in their first 10 years of labor market experience (varying little with bandwidth). In columns (4) to (7) we present specification that allows that the return to being approved at a selective university to change along a worker's career. Under this specification, we estimate a 10%-14% of wage premium for those above the cutoff in their first year of labor market experience, but this differential decreases by 1.5 to 2.7 percentage points per year of experience.

In Panel B of Table 2.11 we present the earnings discontinuity estimates taking into consid-

⁶³Due to space constraint we omit the tests here, but they are available under request.

eration that not all applicants with sufficiently high scores enroll in the top universities. For this purpose, we estimate an earnings equation using a two-stage least square method, where both graduating from a prestigious university and its interaction with experience are instrumented with an indicator for PAA scores above the admission cutoff and its interaction with experience. We estimate a 16-22% effect of graduating from a selective university on earnings of recent college graduates. However, this gap decreases by 2.1-3.7 percentage points per year of experience in the labor market. Note that these estimates should be interpreted as the casual effect only for those applicants that would enroll in a prestigious university and graduate in the event of achieving a sufficiently high score (intent-to-treat effect).

In order to provide a robustness checks for the main RD findings, we present in table 2.12 estimates for the earnings discontinuity at the admission cutoff and its interaction with experience for different model specifications. Precisely, we show in row (1) that our estimates are not sensitive to the exclusion of controls, which is expected if treatment is random around the admission cutoff. In rows (2) and (3) we test how our estimates change with different specifications for the distance from the admission cutoff. Finally we estimate our preferred model for males and females separately. While we estimate similar coefficients for these two groups, we do not find a significant change in the returns to being approved by a prestigious university with experience for women. We notice however that this result is due to large standard errors that might be explained by the fact that we have a smaller fraction of women in the restricted sample.

2.5 Conclusion

This paper tests whether firms statistically discriminate based on the selectivity of the university attended by workers. We first follow the employer learning statistical discrimination test suggested by Altonji and Pierret (2001) and show that the returns to graduating from an elite university in Chile decreases with experience and that the returns to hard-to-observe ability correlates increase with experience. These results are interpreted as evidence for statistical discrimination based on university selectivity.

Furthermore, we take advantage of the centralized admission process of traditional universities in Chile to propose a statistical discrimination test based on a regression discontinuity design. We show that recent graduates just above the admission cutoff to the most prestigious universities in Chile have significantly higher earnings than those just below the cutoff. However, as workers gain labor market experience, the earnings gap between these two groups decreases to the point that we cannot reject zero wage differentials 4 years after graduation. We interpret this result as firms paying workers in accordance with the selectivity of their college when they graduate from school, but rewarding them based on their true productivity as they reveal their quality to employers.

Our results shed some light on the benefits of graduating from a selective university. We interpret our findings as evidence that attending a prestigious university has a significant impact on signaling to firms a worker's unobservable quality. However, employers learn fast and individuals tend to be paid in accordance with their true ability as they gain experience in the labor market.

Table 2.9: Descriptive Statistics for Selective and Non-Selective Universities

Variables	Selective Universities		Non-selective Universities	
	Mean	Std. Dev.	Mean	Std.Dev.
Female	0.53	0.5	0.56	0.5
Language PAA Score	680.6	61.3	588.8	82.8
Math PAA Score	715.9	68.6	609.4	103.4
High School Grade	644.5	78.7	575.1	94.1
Private High School	0.11	0.32	0.07	0.25
Number of Individuals	11,554		46,923	

Note: Math and Language PAA scores are components of the centralized test for admission in University in Chile. See section 3 for definition of selective university.

Table 2.10: Earnings for Selective and Non-Selective Universities

Statistic	Selective Universities	Non-selective Universities
Log of Annual Wage (in 1999 Pesos)		
Mean	15.58	15.19
Std. Deviation	1.09	1.16
Observations	61,844	251,233

Note: See section 3 for definition of selective university.

Table 2.11: EL-SD Regression Discontinuity Test

Dependent Variable: Log Annual Wage						
Panel A	Reduced Form					
	Bandwidth (Points from Cutoff)			Bandwidth (Points from Cutoff)		
	125	100	75	125	100	75
Model	(1)	(2)	(3)	(4)	(5)	(6)
Approved at Selective University	0.081 (0.0287)***	0.073 (0.0308)**	0.065 (0.0342)*	0.143 (0.0330)***	0.119 (0.0348)***	0.099 (0.0382)***
Approved at Selective Univ.* Experience				-0.027 (0.0071)***	-0.020 (0.0073)***	-0.015 (0.0078)*
Observations	39,748	36,639	31,843	39,748	36,639	31,843
R-squared	0.135	0.128	0.120	0.135	0.128	0.120
Panel B	2 Stages Least Square					
	Bandwidth (Points from Cutoff)			Bandwidth (Points from Cutoff)		
	125	100	75	125	100	75
Model	(1)	(2)	(3)	(4)	(5)	(6)
Graduated from Selective University	0.133 (0.0470)***	0.122 (0.0512)**	0.110 (0.0576)*	0.222 (0.0525)***	0.189 (0.0565)***	0.160 (0.0629)**
Graduated from Selective Univ.* Experience				-0.037 (0.0098)***	-0.028 (0.010)***	-0.021 (0.0109)*
Observations	39,748	36,639	31,843	39,748	36,639	31,843
R-squared	0.140	0.133	0.125	0.140	0.134	0.126

Approved at Selective Univ.: Points from the Cutoff \geq 0

Controls: Points from the Cutoff, and Interaction of Points from the Cutoff with Approved at Prestigious Univ., Female, Cubic Experience Polynomial, Major Dummies, and Year Dummies.

Instrument in Panel B: In columns (1)-(6) the endogenous variables are instrumented with Approved at Prestigious University and in columns (4)-(6) also with its interaction with experience

White/Huber standard errors accounting clustered at the individual level are reported in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: The sample is restricted to individuals with engineering, business, medical and law degree (see section 5 for details).

Table 2.12: EL-SD Regression Discontinuity Test - Robustness Checks

Dependent Variable: Log Annual Wage, Applicants within 100 points from Cutoff

Regression Specification	Additional Controls	Function of Points from the Cutoff	Flexible Coefficient?	Sample	Estimated Coefficients	
					Approved at Selective University	Approved at Selective Univ.* Experience
(1)	No	Linear	Yes	All	0.130 (0.0354)***	-0.019 (0.0075)**
(2)	Yes	Cubic	No	All	0.103 (0.0426)**	-0.020 (0.0073)***
(3)	Yes	Cubic	Yes	All	0.109 (0.0425)**	-0.020 (0.0073)***
(4)	Yes	Linear	Yes	Males	0.100 (0.0427)**	-0.018 (0.0086)**
(5)	Yes	Linear	Yes	Females	0.153 (0.0599)**	-0.022 (0.0135)

All specifications include Cubic Experience Polynomial.

Approved at Selective Univ.: Points from the Cutoff \geq 0

Additional Controls: Female, Major Dummies, and Year Dummies.

Flexible coefficient indicates whether the estimated coefficients of points from cutoff was allowed to differ on each side of the admission cutoff
White/Huber standard errors clustered at the individual level are reported in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: The sample is restricted to individuals with engineering, business, medical and law degree (see section 5 for details).

Figure 2.12: Application Process to Traditional Universities

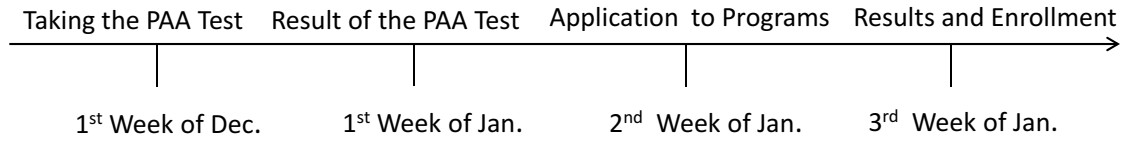
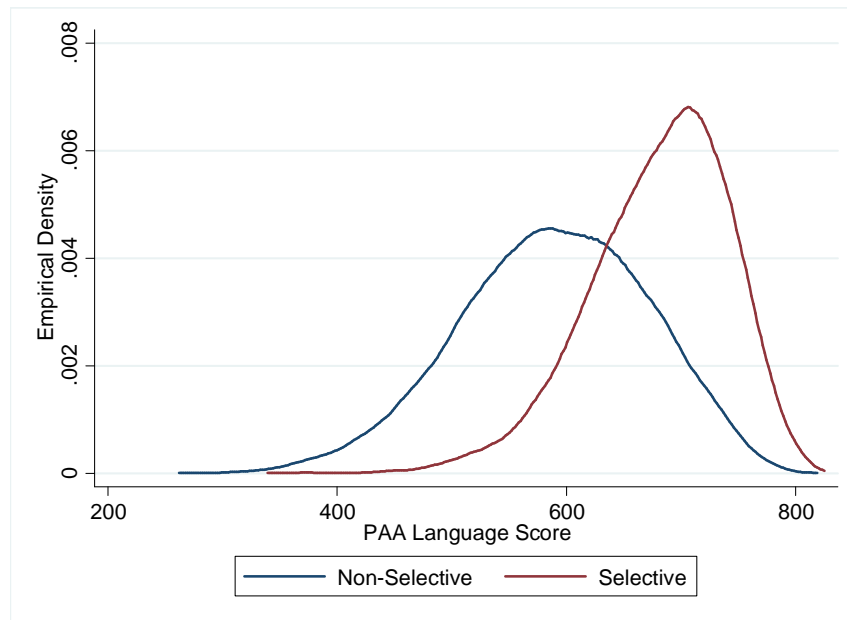
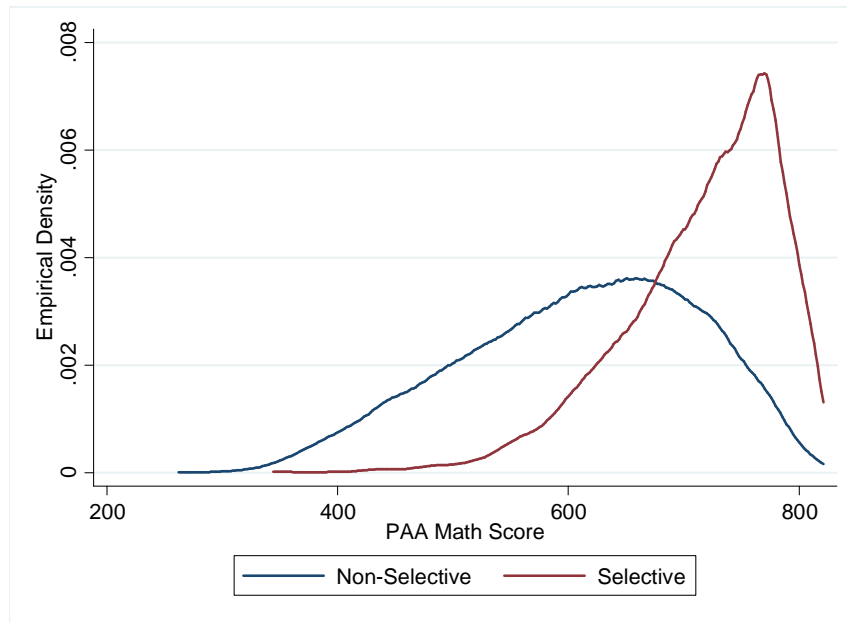


Figure 2.13: Smoothed Language PAA Score Distribution



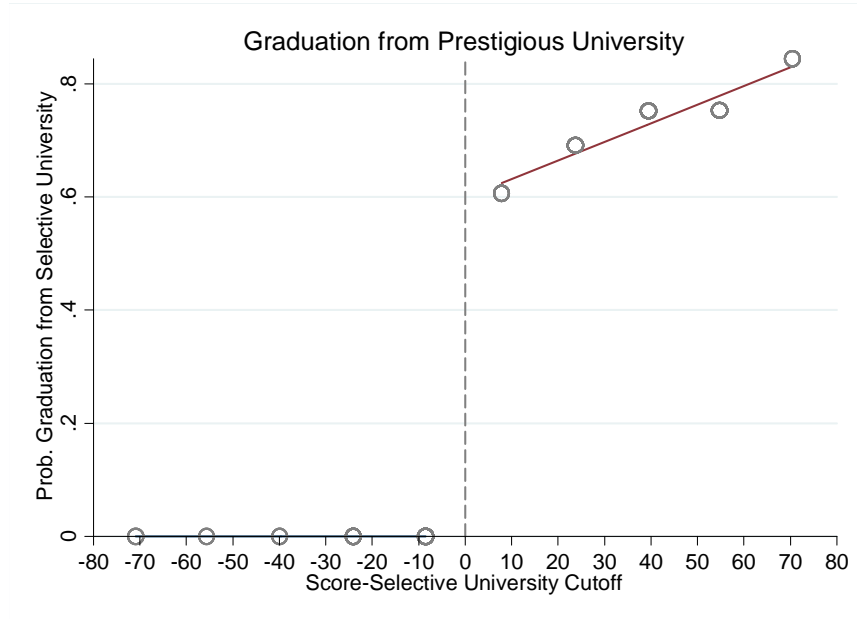
Note: Language PAA is a component of the centralized test for admission to university in Chile. See section 3 for definition of selective university.

Figure 2.14: Smoothed Math PAA Score Distribution



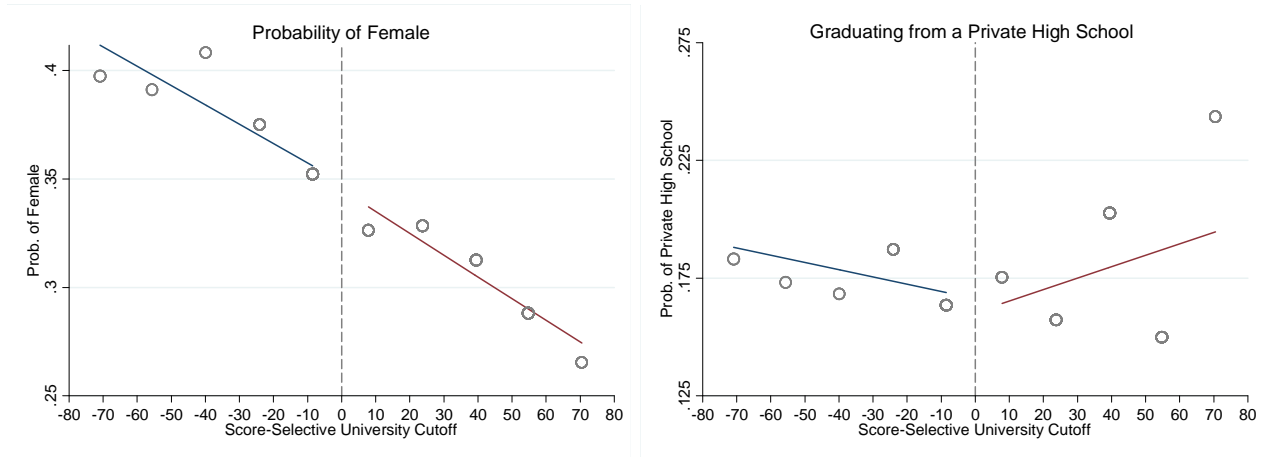
Note: Math PAA is a component of the centralized test for admission to university in Chile. See section 3 for definition of selective university.

Figure 2.15: Graduation from Selective University Discontinuity



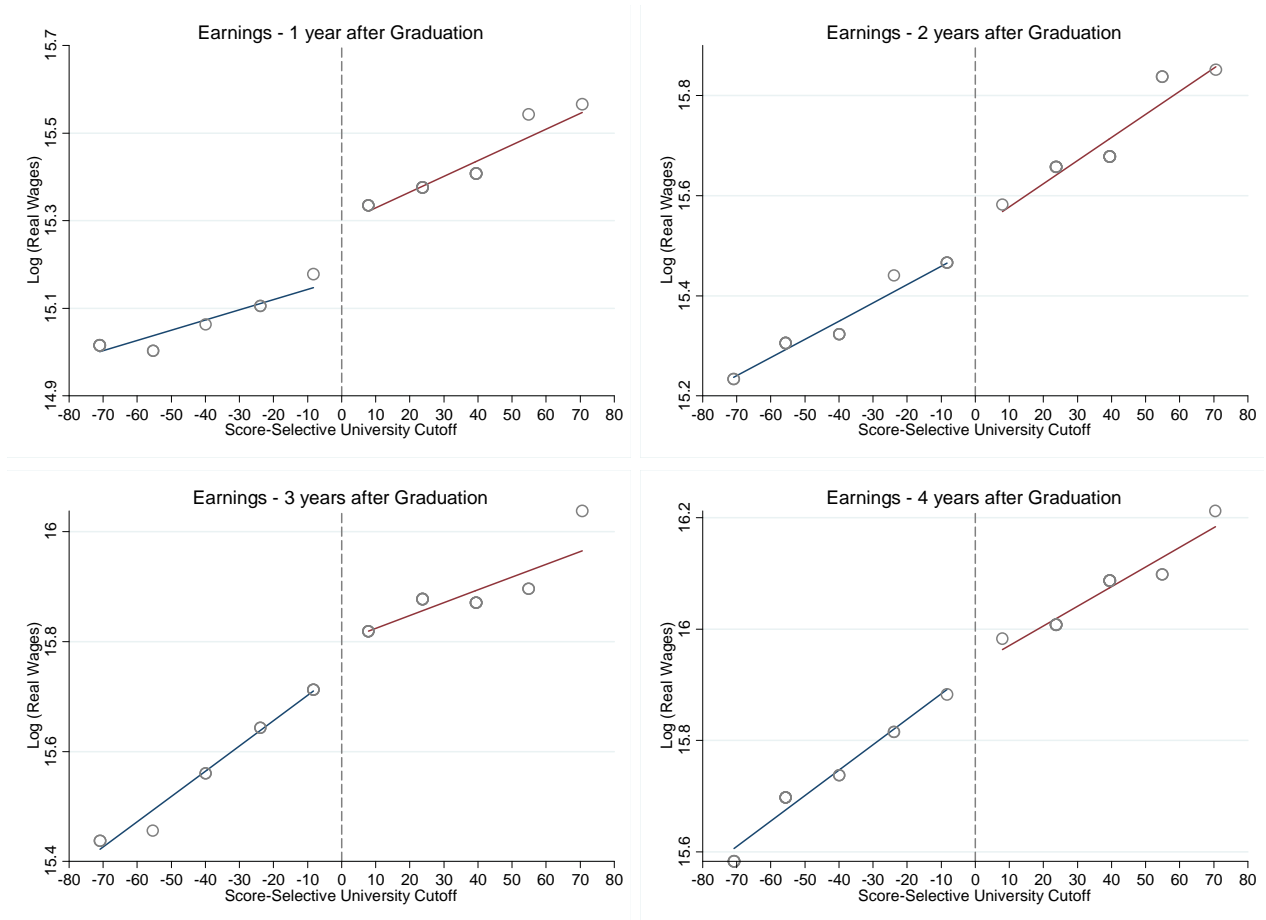
Note: Open circles represent 16 points local averages and the lines are local linear fits below and above the admission cutoff. The sample is restricted to individuals with engineering, business, medical and law degrees (see section 5 for details).

Figure 2.16: Discontinuity at Pre-treatment Outcomes



Note: Open circles represent 16 points local averages and the lines are local linear fits below and above the admission cutoff. The sample is restricted to individuals with engineering, business, medical and law degrees (see section 5 for details).

Figure 2.17: Earnings Discontinuity by Experience



Note: Earnings are defined as log annual wages measured in real Chilean pesos. Open circles represent 16 points local averages and the lines are local linear fits below and above the admission cutoff. The sample is restricted to individuals with engineering, business, medical and law degree (see section 5 for details).

CHAPTER III

Recruitment of Foreigners in the Market for Computer Scientists in the US⁶⁴

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

⁶⁴This chapter was written with John Bound, Joseph M. Golden and Gaurav Khanna.

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.⁶⁵

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

⁶⁵In 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.

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.⁶⁶ 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 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.

⁶⁶These 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.

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 (2004) 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 States⁶⁷ and a substantial increase in the number of Internet users. One 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%.⁶⁸ Table 3.13 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 3.18a, we use

⁶⁷The 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)).

⁶⁸Here 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.

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 3.18b 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 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 3.18c 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 1998.⁶⁹

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

Employment adjustments in the market for computer scientists happened disproportionately

⁶⁹It 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.

among foreigners during the Internet boom. Evidence for this claim is found in Table 3.13 and Figure 3.18d, where we use the Census and CPS to compare the share of foreign computer scientists to the share of foreign workers in other occupations.⁷⁰ 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 Act of 1990 established the H-1B visa program for temporary workers in “specialty occupations.”⁷¹ 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,

⁷⁰Here 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.

⁷¹The 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.

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 for FY1999 and to 195,000 for FY2000-2003. The cap then reverted to 65,000.⁷² Figure 3.19 shows the growth in the number of H-1 visas issued

⁷²The 2000 legislation that raised the cap also excluded Universities and non-profit research facilities from

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.

it, and a 2004 change added an extra 20,000 visas for foreigners who received a masters degree in the US

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 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,⁷³ this paper proposes an alternative interpretation 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

⁷³If computer scientists have large effects on firm productivity, then demand curves for them would be very elastic. 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.

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 computer 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.⁷⁴ 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 us to distinguish visa types.

⁷⁴Here 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).

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.

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.⁷⁵

⁷⁵Essentially, 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

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 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 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.

of college graduates with age a in time period t ,⁷⁶ the number of recent graduates with a computer science degree in year t is represented by $C_t = p_t^c M_t^{22}$.

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 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 (2009)). 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.⁷⁷

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:

$$V_{t,a}^c = \max\{w_t^c + \beta \mathbb{E}_t V_{t+1,a+1}^c + \varepsilon_{it}^c, w_t^o - c(a) + \beta \mathbb{E}_t V_{t+1,a+1}^o + \varepsilon_{it}^o + \alpha_1\}$$

⁷⁶We 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

⁷⁷The 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.

$$V_{t,a}^o = \max\{w_t^c - c(a) + \beta\mathbb{E}_t V_{t+1,a+1}^c + \varepsilon_{it}^c, w_t^o + \beta\mathbb{E}_t V_{t+1,a+1}^o + \varepsilon_{it}^o + \alpha_1\}$$

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_t^o is exogenous and perfectly anticipated by the workers.⁷⁸ 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 independently⁷⁹ and identically distributed and for $s = \{c, o\}$ can be defined as $\varepsilon_{it}^s = \sigma_1 v_{it}^s$ where σ_1 is a 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

⁷⁸As a matter of fact, in the simulations of the paper we will set $w_t^o = 1$ and measure wages of computer scientists as an occupational premium.

⁷⁹In 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)).

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:

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

(3.27)

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 don't 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 correspondent 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.⁸⁰

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

$$L_t = \sum_{a=22}^{a=65} L_t^a \quad (3.28)$$

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.

⁸⁰In 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.

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 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).⁸¹ 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.⁸²

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.⁸³ 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:

$$F_t = \sum_{j=0}^5 R_{t-j} \quad (3.29)$$

⁸¹As 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.

⁸²In 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.

⁸³The 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.

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 input used in the production function and we ignore the firm's decision about capital or other types of labor adjustments.⁸⁴ 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).⁸⁵ 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,

⁸⁴The 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)).

⁸⁵While we suspect it would make sense to allow workers of different ages to be imperfect substitutes in production for each other, CPS sample sizes are too small to support this kind of analysis.

there are no adjustment costs for American workers but firms incur extra costs to recruit foreigners.⁸⁶ 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.

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:⁸⁷

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

subject to foreign labor supply:

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

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

⁸⁶In 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.

⁸⁷For 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.

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, because increasing marginal recruitment costs of foreigners, prevents firms from

completely substituting foreigners for domestic workers.⁸⁸

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:

$$A_t \gamma (L_t + \theta F_t)^{\gamma-1} = w_t \quad (3.30)$$

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:

$$\theta A_t \gamma (L_t + \theta F_t)^{\gamma-1} - w_t - c_1 - 2c_2 R_t + \sum_{j=1}^5 \beta^j \mathbb{E}_t \left[\frac{\partial \pi_{t+j}}{\partial R_t} \right] = 0$$

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:

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

⁸⁸Our 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.

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

$$\sum_{j=0}^5 \beta^j \mathbb{E}_t [\theta A_{t+j} \gamma (L_{t+j} + \theta F_{t+j})^{\gamma-1} - w_{t+j}] = c_1 + 2c_2 R_t \quad (3.31)$$

Equation (3.31) 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 (3.27) characterizes the expectations of workers with respect to future career prospects, equations (3.28) and (3.29) are the dynamic labor supply of American and foreigner computer scientists respectively, and equations (3.30) and (3.31) 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,⁸⁹ such that:

⁸⁹We model the technology progress as a close to random walk since we will interpret the Internet boom as a

$$A_t = 0.999A_{t-1} + 0.001\bar{A} + \xi_t \quad (3.32)$$

where \bar{A} is the steady state level of progress, and ξ_t is the i.i.d. random idiosyncratic productivity shock with mean zero that is assumed to be independent of other variables of the model.⁹⁰

The equilibrium of the model can be expressed by a mapping from the state variables: $s = \{C_t, L_{t-1}^{22}, \dots, L_{t-1}^{64}, R_{t-1}, \dots, R_{t-5}, A_{t-1}\}$ and exogenous productivity shock ξ_t to the values of L_t , w_t , R_t , and $\mathbf{V}t$, the vector of career prospects at different occupations for different ages, that satisfies the system of equations (3.27) to (3.32). 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).

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.

⁹⁰Note 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.

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 productivity⁹¹ parameter $\theta = 1.12$ based on estimations from the 2003 National Survey of College Graduates data.⁹² 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 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 enrollment⁹³ between 1994 and 2004. We define other STEM occupations

⁹¹In 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.

⁹²Specifically, 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.

⁹³Uses data from 1996 to 2006, representing enrollment decisions from 1994 to 2004. See the Online Appendix for more details.

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 wages⁹⁴ of other STEM jobs, and thus define our key data series as:⁹⁵

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

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

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

⁹⁴We 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.

⁹⁵See 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: $X_{1994,smooth} = \frac{2}{3}X_{t,raw} + \frac{1}{3}X_{t+1,raw}$.

are surprised by changes in A_t .⁹⁶ Note that following equation (3.32), 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 2004.⁹⁷ We use a Nelder-Mead simplex method to find parameter values which yield solutions to the model under these data restrictions.⁹⁸

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 .

⁹⁶We 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.

⁹⁷The decision to match changes in the age structure of CS rather than 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.

⁹⁸Note 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_{t-2}^c in two years and the observation of changes in $s_t^{22,34}$, $s_t^{35,44}$, and $s_t^{45,65}$ in 2004.

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}.⁹⁹ We present the calibrated parameters for these different values of γ in Table 3.14 and a comparison of the data with the model's output in Figures 3.20 - 3.21. 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 3.14. 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 3.14, 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 11.6.¹⁰⁰ 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

⁹⁹ γ in the 0.25 to 0.75 range imply labor demand elasticities between -1.33 and -4.0. Ryoo and Rosen (2004), 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.

¹⁰⁰Ryoo 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.

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 3.20-3.21, 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 our model has trouble fitting the data in this period for some series. Figure 3.20 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 3.21 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 3.20-3.21 (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 3.15 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 γ '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,¹⁰¹ 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

¹⁰¹According 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.

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.

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 γ '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 (2004). 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 3.13: 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 Bachelor's/Master's	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 3.14: 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
$\epsilon_{L_s,w}^s$	Short run US occupational labor supply elasticity	0.5591	0.6745	0.6642
$\epsilon_{p,w}^s$	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.

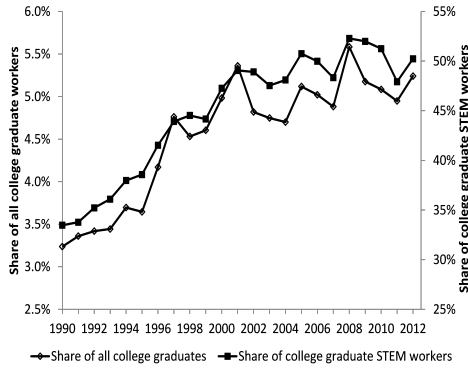
Table 3.15: Summary of Results from Counterfactual Simulation

% 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%

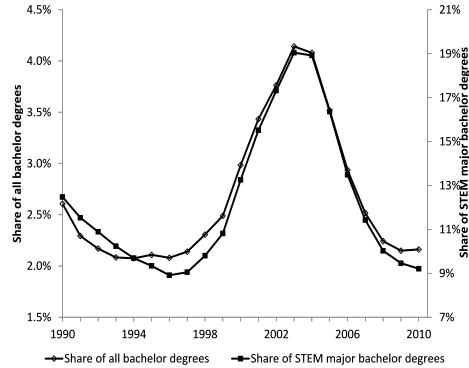
Note: The counterfactual simulates an economy from 1994-2009 in which the level of foreign CS workers is not allowed to increase from its 1994 value. The parameter γ determines the labor demand elasticity to wages. See section 4 for details.

Figure 3.18: Major Trends (1990 to 2012)

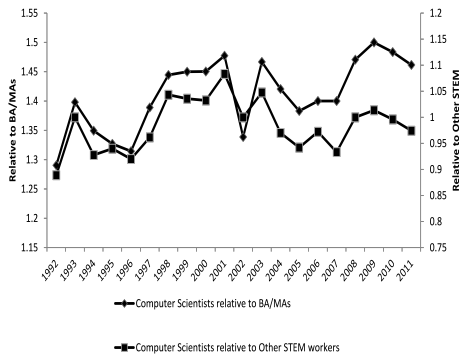
(a) Fraction of Computer Scientists in US Workforce



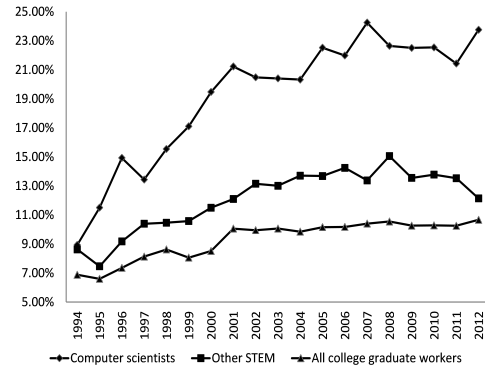
(b) Computer Science Fraction of Bachelor Degrees Awarded in US



(c) Relative Earnings of Computer Scientists



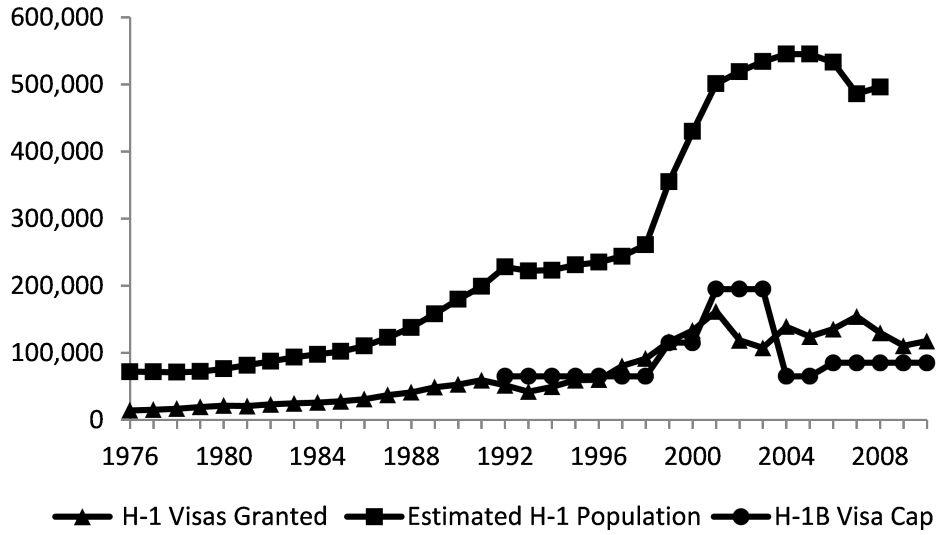
(d) Foreign Born and Immigrated at Age 18 or Older Fraction of Employed Population by Occupation



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.

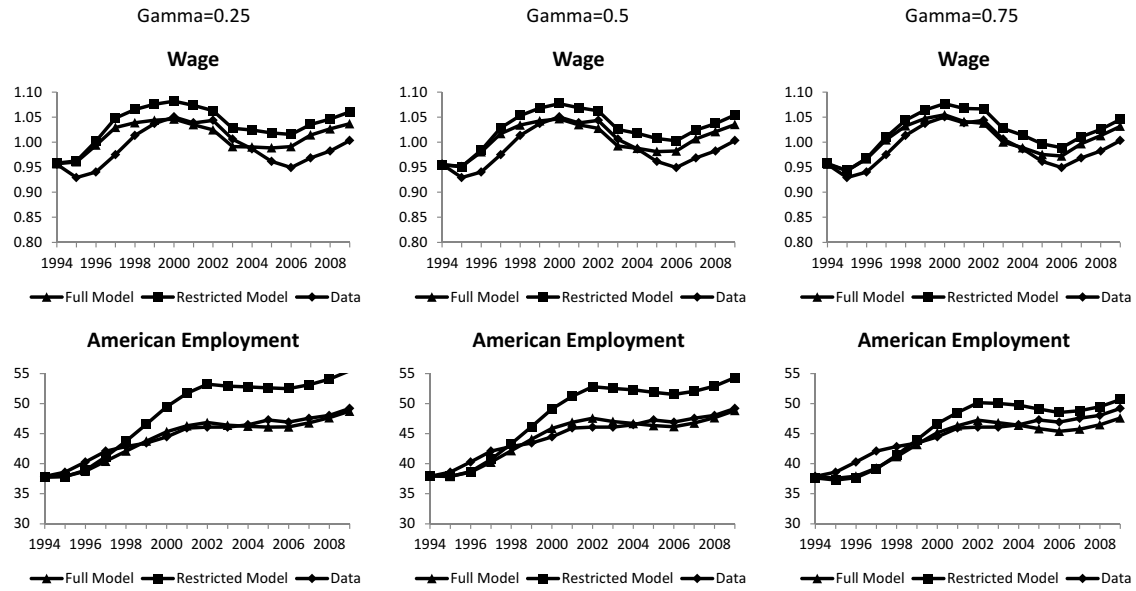
Sources: March CPS (for employment, earnings, and immigrants); IPEDS Completions Survey (for degrees)

Figure 3.19: H-1 and H-1B Visa Population



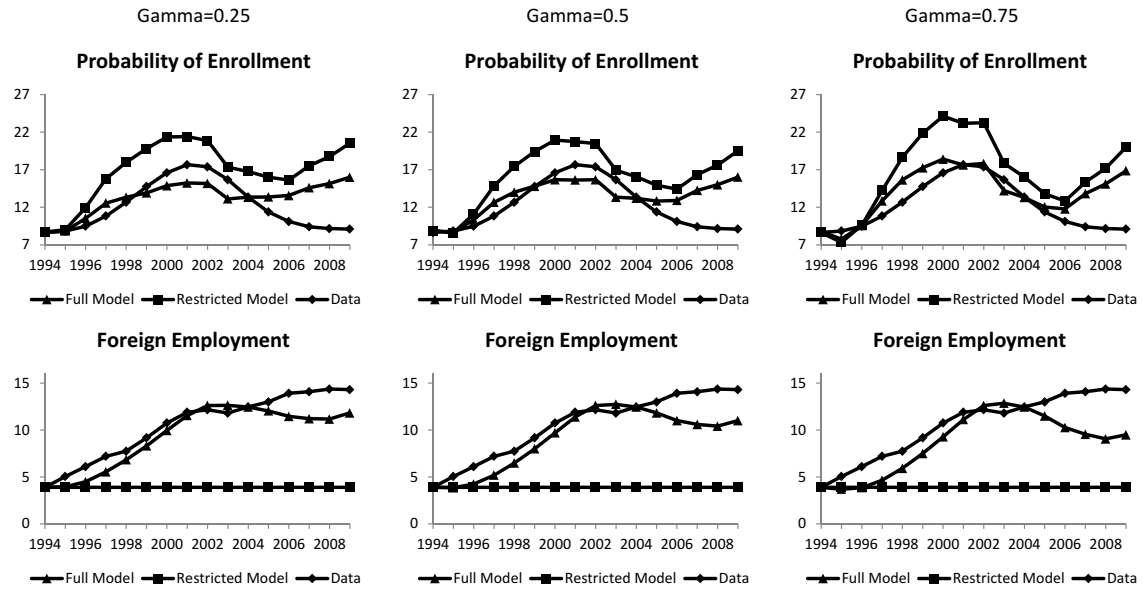
Note: Population stock is constructed using estimations of inflow (visas granted) and outflow (deaths, permanent residency, or emigration) of H-1 workers. In later years, the number of visas granted could exceed the visa cap due to exemptions for foreigners who work at universities and non-profit research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

Figure 3.20: 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 γ determines the labor demand elasticity to wages. See Section 4 for details.

Figure 3.21: 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 γ determines the labor demand elasticity to wages. See Section 4 for details.

Bibliography

- Acemoglu, D., Pischke, J. S., 1998. Why do firms train? Theory and evidence. *The Quarterly Journal of Economics* 113 (1), pp. 79–119.
- Aigner, D. J., Cain, G. G., 1977. Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review* 30 (2), pp. 175–187.
- Albrecht, J. W., Edin, P.-A., Sundstrom, M., Vroman, S. B., 1999. Career interruptions and subsequent earnings: A reexamination using Swedish data. *The Journal of Human Resources* 34 (2), pp. 294–311.
- Altonji, J. G., Blank, R. M., 1999. Race and gender in the labor market. *Handbook of Labor Economics* 3, 3143–3259.
- Altonji, J. G., Pierret, C. R., February 2001. Employer learning and statistical discrimination. *The Quarterly Journal of Economics* 116 (1), 313–350.
- Antecol, H., Bedard, K., 2004. The racial wage gap: The importance of labor force attachment differences across black, mexican, and white men. *The Journal of Human Resources* 39 (2), pp. 564–583.
- Arcidiacono, P., Bayer, P., Hizmo, A., 2010. Beyond signaling and human capital: Education

- and the revelation of ability. *American Economic Journal: Applied Economics* 2 (4), 76–104.
- Arrow, K. J., 1962. The economic implications of learning by doing. *The Review of Economic Studies* Vol 29, Issue 3, 155–173.
- Arulampalam, W., 2001. Is unemployment really scarring? Effects of unemployment experiences on wages. *The Economic Journal* 111 (475), pp. F585–F606.
- Becker, G. S., 1962. Investment in human capital: A theoretical analysis. *Journal of Political Economy* 70 (5), pp. 9–49.
- Ben-Porath, Y., 1967. The production of human capital and the life cycle of earnings. *The Journal of Political Economy* 75 (4), 352–365.
- Black, D. A., Smith, J. A., 2006. Estimating the returns to college quality with multiple proxies for quality. *Journal of Labor Economics* 24 (3), pp. 701–728.
- Blackburn, M. L., Neumark, D., 1995. Are OLS estimates of the return to schooling biased downward? Another look. *The Review of Economics and Statistics* 77 (2), pp. 217–230.
- Blau, F. D., Kahn, L. M., 2013. The feasibility and importance of adding measures of actual experience to cross-sectional data collection. *Journal of Labor Economics* 31 (2), pp. S17–S58.
- Bollinger, C., Hirsch, B., 2007. 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.

- Borjas, G., 1999. The economic analysis of immigration. Handbook of Labor Economics Vol 3A, edited by Orley Ashenfelter and David Card, North-Holland.
- Borjas, G., 2009. 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 Richard B. Freeman and Daniel L. Goroff, eds., 131 – 161.
- Bound, J., Braga, B., Golden, J., Turner, S., 2013. Pathways to adjustment: The case of information technology workers. American Economic Review: Papers and Proceedings.
- Brewer, D. J., Eide, E. R., Ehrenberg, R. G., 1999. Does it pay to attend an elite private college? Cross-cohort evidence on the effects of college type on earnings. Journal of Human Resources, 104–123.
- Clemens, M. A., 2013. Why do programmers earn more in houston than hyderabad? evidence from randomized processing of us visas. American Economic Review, Papers & Proceedings 103(3): 198-202.
- Collard, F., Juillard, M., 2001a. A higher-order taylor expansion approach to simulation of stochastic forward-looking models with an application to a non-linear phillips curve. Computational Economics 17(2), 125–139.
- Collard, F., Juillard, M., 2001b. Accuracy of stochastic perturbation methods: The case of asset pricing models. Journal of Economic Dynamics and Control 25(7), 979–999.
- Corcoran, M., Duncan, G. J., 1979. Work history, labor force attachment, and earnings differences between the races and sexes. Journal of Human Resources, 3–20.

- Dale, S. B., Krueger, A. B., November 2002. Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *The Quarterly Journal of Economics* 117 (4), 1491–1527.
- DeVaro, J., Waldman, M., 2012. The signaling role of promotions: Further theory and empirical evidence. *Journal of Labor Economics* 30 (1), pp. 91–147.
- Farber, H. S., Gibbons, R., 1996. Learning and wage dynamics. *The Quarterly Journal of Economics* 111 (4), pp. 1007–1047.
- Filer, R. K., 1993. The usefulness of predicted values for prior work experience in analyzing labor market outcomes for women. *The Journal of Human Resources* 28 (3), pp. 519–537.
- Freeman, R. B., March 1975. Supply and salary adjustments to the changing science manpower market: Physics, 1948-1973. *American Economic Review* 65 (1), 27–39.
- Freeman, R. B., 1976. A cobweb model of the supply and starting salary of new engineers. *Industrial and Labor Relations Review* 29 (2), pp. 236–248.
- Freeman, R. B., 2006. Does globalization of the scientific/engineering workforce threaten u.s. economic leadership? *Innovation Policy and the Economy* Volume 6, Adam B. Jaffe, Josh Lerner, and Scott Stern.
- Freeman, R. B., May 2009. What does global expansion of higher education mean for the us? Working Paper 14962, National Bureau of Economic Research.
- Gibbons, R., Katz, L. F., October 1991. Layoffs and lemons. *Journal of Labor Economics* 9 (4), 351–80.

- Gibbons, R., Waldman, M., 2006. Enriching a theory of wage and promotion dynamics inside firms. *Journal of Labor Economics* 24 (1), pp. 59–107.
- Hamermesh, D. S., 1989. Labor demand and the structure of adjustment costs. *The American Economic Review* 79 (4), pp. 674–689.
- Hastings, J. S., Neilson, C. A., Zimmerman, S. D., 2013. Are some degrees worth more than others? Evidence from college admission cutoffs in chile. Tech. rep., National Bureau of Economic Research.
- Heckman, J., Lochner, L., Todd, P., 2006. Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. *Handbook of the Economics of Education* 1, 307–458.
- Hershbein, B., 2013. Worker signals among new college graduates: The role of selectivity and GPA. Working Paper, Upjohn Institute.
- Hoekstra, M., 2009. The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The Review of Economics and Statistics* 91 (4), 717–724.
- Hoxby, C. M., 1998. The return to attending a more selective college: 1960 to the present. Unpublished manuscript, Department of Economics, Harvard University, Cambridge, MA.
- Hu, L., Taber, C., 2011. Displacement, asymmetric information, and heterogeneous human capital. *Journal of Labor Economics* 29 (1), 113–152.
- Hunt, J., Gauthier-Loiselle, M., 2010. How much does immigration boost innovation? *American Economic Journal: Macroeconomics* 2(2), 31–56.

- Imbens, G. W., Lemieux, T., 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142 (2), 615–635.
- Juillard, M., 1996. Dynare: A program for the resolution and simulation of dynamic models with forward variables through the use of a relaxation algorithm. CEPREMAP 9602.
- Kambourov, G., Manovskii, I., 2009. Occupational specificity of human capital. *International Economic Review* 50 (1), 63–115.
- Kaufmann, K. M., Messner, M., Solis, A., 2012. Returns to elite higher education in the marriage market: Evidence from Chile. Tech. rep., Bocconi University working paper.
- Kerr, S. P., Kerr, W. R., Lincoln, W. F., 2013. Skilled immigration and the employment structures of u.s. firms. Working Paper.
- Kerr, W., Lincoln, W., 2010. The supply side of innovation: H-1b visa reforms and u.s. ethnic invention. *Journal of Labor Economics* 28(3), 473–508.
- Kim, M.-K., Polachek, S. W., 1994. Panel estimates of male-female earnings functions. *Journal of Human Resources*, 406–428.
- Kirkegaard, J., 2005. 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.
- Kline, P., 2008. 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.

- Kroft, K., Lange, F., Notowidigdo, M. J., 2013. Duration dependence and labor market conditions: Evidence from a field experiment. *The Quarterly Journal of Economics*.
- Krugman, P., 1979. A model of innovation, technology transfer, and the world distribution of income. *Journal of Political Economy* Vol. 87, No. 2, 253–266.
- Lang, K., Siniver, E., 2011. Why is an elite undergraduate education valuable? evidence from israel. *Labour Economics* 18 (6), 767 – 777.
- Lange, F., 2007. The speed of employer learning. *Journal of Labor Economics* 25 (1), 1–35.
- Lee, D. S., Lemieux, T., September 2010. Regression discontinuity designs in economics. *Journal of Economic Literature* 48 (2), 281–355.
- Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., Postel, J., Roberts, L. G., Wolff, S. S., Feb. 1997. The past and future history of the internet. *Commun. ACM* 40 (2), 102–108.
- Lemieux, T., 2006. The Mincer equation thirty years after schooling, experience, and earnings. In: Grossbard-Shechtman, S. (Ed.), *Jacob Mincer, A Pioneer of Modern Labor Economics*. Springer Verlag.
- Light, A., Ureta, M., January 1995. Early-career work experience and gender wage differentials. *Journal of Labor Economics* 13 (1), 121–54.
- Lockwood, B., 1991. Information externalities in the labour market and the duration of unemployment. *The Review of Economic Studies* 58 (4), 733–753.
- Mansour, H., 2012. Does employer learning vary by occupation? *Journal of Labor Economics* 30 (2), 415–444.

- Matloff, N., 2003. On the need for reform of the h-1b non-immigrant work visa in computer-related occupations. *University of Michigan Journal of Law Reform* 36, no. 4, 815–914.
- Mincer, J., 1962. On-the-job training: Costs, returns, and some implications. *The Journal of Political Economy* 70 (5), 50–79.
- Mincer, J., Ofek, H., 1982. Interrupted work careers: Depreciation and restoration of human capital. *The Journal of Human Resources* 17 (1), pp. 3–24.
- Mincer, J., Polachek, S., 1974. Family investments in human capital: Earnings of women. *Journal of Political Economy* 82 (2), pp. S76–S108.
- Mincer, J. A., May 1974. Schooling, Experience, and Earnings. No. minc74-1 in NBER Books. National Bureau of Economic Research, Inc.
- Mithas, S., Lucas, H. C., May 2010. Are foreign it workers cheaper? u.s. visa policies and compensation of information technology professionals. *Management Science* 56 (5), 745–765.
- Mukhopadhyay, S., Oxborrow, D., 2012. The value of an employment-based green card. *Demography* 49, 219–237.
- Murphy, K. M., Welch, F., 1990. Empirical age-earnings profiles. *Journal of Labor Economics* 8 (2), pp. 202–229.
- Murphy, K. M., Welch, F., 1992. The structure of wages. *The Quarterly Journal of Economics* 107 (1), 285–326.
- Nathan, M., 2013. The wider economic impacts of high-skilled migrants: A survey of the literature. IZA Working Papers Discussion Paper No. 7653.

- Peri, G., Shih, K., Sparber, C., 2013. Stem workers, h1b visas and productivity in us cities.
- Peri, G., Sparber, C., 2011. Highly-educated immigrants and native occupational choice. *Industrial Relations* Vol. 50 (3), 385–411.
- Phelps, E. S., 1972. The statistical theory of racism and sexism. *The American Economic Review* 62 (4), pp. 659–661.
- Pop-Eleches, C., Urquiola, M., 2013. Going to a better school: Effects and behavioral responses. *The American Economic Review* 103 (4), 1289–1324.
- Regan, T., Oaxaca, R., April 2009. Work experience as a source of specification error in earnings models: implications for gender wage decompositions. *Journal of Population Economics* 22 (2), 463–499.
- Romer, P. M., 1986. Increasing returns and long-run growth. *Journal of Political Economy* Vol. 94, No. 5, 1002–1037.
- Rust, J., 1987. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica* 55 (5), pp. 999–1033.
- Ryoo, J., Rosen, S., February 2004. The engineering labor market. *Journal of Political Economy* 112 (S1), S110–S140.
- Saavedra, J. E., 2008. The returns to college quality: A regression discontinuity analysis. Unpublished manuscript, Harvard University, Cambridge, MA.
- Sargent, T. J., 1978. Estimation of dynamic labor demand schedules under rational expectations. *Journal of Political Economy* 86 (6), pp. 1009–1044.

- Schönberg, U., 2007. Testing for asymmetric employer learning. *Journal of Labor Economics* 25 (4), 651–691.
- Shapiro, M. D., 1986. The dynamic demand for capital and labor. *The Quarterly Journal of Economics* 101 (3), pp. 513–542.
- Spence, M., 1973. Job market signaling. *The Quarterly Journal of Economics* 87 (3), pp. 355–374.
- U.S. Department of Commerce, 2000. Digital economy 2000. Tech. rep., U.S. Department of Commerce.
- U.S. General Accounting Office, 1992. Immigration and the labor market: Nonimmigrant alien workers in the united states. Tech. Rep. GAO/PEMD-92-17., U.S. General Accounting Office, Washington, D.C.
- U.S. General Accounting Office, 2000. H-1b foreign workers: Better controls needed to help employers and protect workers. Tech. Rep. GAO/HEHS 00-157., U.S. General Accounting Office, Washington, D.C.
- U.S. General Accounting Office, 2011. H-1b visa program: Reforms are needed to minimize the risks and costs of current program. Tech. Rep. GAO/HEHS 00-157., U.S. General Accounting Office, Washington, D.C.
- U.S. Immigration and Naturalization Service, February 2000. Characteristics of specialty occupation workers (h 1b). Tech. rep., U.S. Immigration and Naturalization Service, Washington, D.C.

Yatchew, A., 1998. Nonparametric regression techniques in economics. *Journal of Economic Literature* 36 (2), 669–721.