ESSAYS IN LABOR ECONOMICS

By

Bryce VanderBerg

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Economics – Doctor of Philosophy

2022

ABSTRACT

ESSAYS IN LABOR ECONOMICS

By

Bryce VanderBerg

This dissertation consists of two empirical studies and one applied theoretical study in labor economics. In the first chapter, I study the extend to which an observed layoff is used by employers to infer a worker's unobserved ability early in their labor market career. In the second chapter, I develop a theoretical model of wage dynamics that extends the employer learning and statistical discrimination model of Altonji and Pierret (2001) to allow for discrete changes in observable characteristics. In the third chapter, which is joint work with Gabrielle Pepin at the W.E. Upjohn Institute, we study the contribution of occupational sorting and mismatch to child penalties in the United States.

I: The Signaling Role of Early Career Job Loss

I examine the extent to which ability signaling explains long-term wage losses suffered by young workers who experience layoffs. Young workers are of particular interest because employers have limited information about their ability, so signaling theoretically plays a larger role in determining wages. In addition, young workers are unlikely to experience wage losses due to loss of industry-specific human capital or separation from high-quality job matches, which may explain long-term wage decreases among older workers. Using data from the National Longitudinal Survey of Youth 1997, I show that young workers of all ability levels initially experience similar wage losses following layoffs, but high-relative ability workers fully recover within five years while low-relative ability workers experience persistent wage losses. Consistent with traditional learning models, relative, not actual, ability affects wage trajectories. I illustrate a conceptual model of layoff signaling that varies by pre-layoff experience and can explain divergent wage trajectories across high- and low-relative ability workers. I test the model empirically and find that low-relative ability workers' inability to overcome negative layoff signals explains a substantial proportion of long-term wage losses among young workers. Employer learning effects vary by race and gender.

II: Employer Learning and Statistical Discrimination with Unexpected Information

The Employer Learning and Statistical Discrimination (EL-SD) model of Altonji and Pierret (2001) assumes that employers learn about a worker's unobserved ability in a smooth, continuous manner, holding observable characteristics constant. In practice, observable characteristics, such as years of education, often

change discretely over time for many workers. I extend the EL-SD model to allow for changes in observable characteristics to influence an employer's belief about a worker's ability. I show that changes in observable characteristics that are correlated with ability lead to discrete changes in employers' beliefs about the worker's ability, interrupting the smooth, continuous employer learning processes described in the EL-SD model. I further show that this discrete change in employer learning is larger for workers early in their labor market career, with the effect diminishing as labor market experience increases. I then use data from the NLSY97 to empirically test these predictions in the context of the signaling role of returning to school. I find suggestive evidence that returning to school to receive a GED or graduate degree sends a positive ability signal to the labor market, while returning to school to receive an associate or bachelor's degree does not.

III: Occupational Sorting, Multidimensional Skill Mismatch, and the Child Penalty among Working Mothers

We study the extent to which occupational sorting explains child penalties—gender gaps in labor market outcomes due to children—among working parents. Using an event-study approach and data from the National Longitudinal Surveys of Youth (NLSY) 1979 and 1997, we estimate that children generate long-run earnings gaps of over $200 per week among working parents. In the NLSY79, we find that children lead mothers to sort into lower-paying occupations in which employees tend to work fewer hours. We estimate that children increase multidimensional occupation-skill mismatch among working mothers by 0.3 standard deviations, relative both to their own levels of mismatch from before birth and to those of fathers. In the NLSY97, results suggest that improvements in labor market outcomes among fathers in response to children, rather than a worsening of labor market outcomes among mothers, seem to drive child penalties.

This thesis is dedicated to my committee, my family, my fellow graduate students, and, of course, Gabrielle. Thank you for always supporting me.

ACKNOWLEDGEMENTS

I would like to thank Steve Woodbury, Todd Elder, Michael Conlin, Amanda Chuan, Gabrielle Pepin, Stacie Bosley, Patrick Turner, Joseph Marchand, Erina Ytsma, Junjie Guo, seminar participants at Michigan State University, Oakland University, Eastern Michigan University, Western Michigan University, The United States Air Force Academy, the U.S. Census Bureau, the U.S. Food and Drug Administration, the Federal Trade Commission, and the W.E. Upjohn Institute for Employment Research, as well as participants at the MSU-UM-UWO Labour Day Conference for helpful comments and suggestions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

THE SIGNALING ROLE OF EARLY CAREER JOB LOSS

1.1 Introduction

The first decade of a worker's labor market career is an important period of wage growth and job mobility (Topel and Ward 1992).[1] During this period, wages grow by about 60 percent on average, with job mobility accounting for around one-third of the increase (Topel and Ward 1992). Job loss during the first decade of a worker's career can prove detrimental, however, as researchers document large, persistent earnings losses that last for up to a decade or more (Kletzer and Fairlie 2003).[2] While researchers show that loss of firm-specific human capital and high-quality job matches largely explains long-run costs of job loss among older, long-tenured workers (Lachowska, Mas, and Woodbury 2020), these mechanisms are unlikely to apply to younger workers with limited labor market experience.[3] Understanding the mechanism responsible for the earnings losses from early career job loss is crucial for establishing effective policy aimed to help young workers overcome these adverse effects.

In this paper, I examine one possible channel for the long-run costs of job loss among young workers: incomplete information about workers' ability early in their careers. Incomplete information may contribute to young workers' earnings losses if job loss serves as a signal of ability. In particular, given recent evidence that employers learn asymmetrically about a worker's ability over time (Pinkston 2009; L. B. Kahn 2013), distressed firms should choose to lay off their lowest-ability workers first, signaling prospective employers that laid-off workers have lower-than-expected ability (Gibbons and Katz 1991).[4] Given young workers' limited labor market experience, prospective employers likely rely heavily on the negative ability signal in determining wages following job loss. Hence, incomplete information may play a considerable role in

[1]This period is associated with 70-80 percent of lifetime earnings growth (Murphy and Welch 1990). See also Keane and Wolpin (1997), Light and McGarry (1998), Neal (1999), Neumark (2002), Liu (2019), and Forsythe (2019).

[2]Additional studies on early career job loss include Stevens (1997), von Wachter and Bender (2006), Fuji, Shiraishi, and Takayama (2018), and Barnette, Odongo, and Reynolds (2021). Empirical studies on the costs of job loss for older, more established workers include Topel (1990), Jacobson, LaLonde, and Sullivan (1993), and Couch and Placzek (2010); Farber (2015; 2017), to name a few. See also the reviews by Fallick (1993); and Carrington and Fallick (2017).

[3]See also Burdett, Carillo-Tudela, and Coles (2020) and Jarosch (2021) for further evidence on what drives the costs of job loss for older, more established workers. See Carrington and Fallick (2017) for a comprehensive review of the theories proposed for the mechanism driving the long-term costs of job loss.

[4]While the extent to which firms actually have discretion over who they lay off is unclear (Kletzer 1998; Oyer and Schaefer 2011) evidence from von Wachter and Bender (2006) suggests that workers who are laid off early in their careers are negatively selected, implying firms likely have at least some ability to selectively lay off lower ability workers first.

explaining the long-run costs of early-career job loss.

In this paper, I first investigate whether, consistent with employer learning, there exist divergent post-layoff wage paths for workers who are above- versus below-average ability, relative to their peers with similar observable characteristics. I use relative, as opposed to actual, ability levels, as employers likely are most interested in a worker's productivity relative to that of workers with similar skills and education levels. Using an event-study framework, I find that laid-off workers who have an above-average residual AFQT score experience nearly identical initial wage losses as laid-off workers with below-average scores. Within six years, however, above-average laid-off workers' wages nearly fully recover relative to a reference group of continually employed workers, while below-average laid-off workers continue to experience persistent wage losses of around 10 percent.[5]

I then illustrate a layoff signaling framework that accounts for the changing nature of information available to employers over a worker's labor market career. The model treats layoff signaling as a form of statistical discrimination, based on the assumptions that some layoffs are due to the workers having lower-than-expected productivity and that prospective employers are unable to accurately identify the reason behind each layoff. To test the model empirically, I extend the empirical employer learning models of H. Farber and Gibbons (1996) and Altonji and Pierret (2001) to allow for the wage returns to ability to vary not only with experience but also with the timing of a layoff. Using a sample of young workers from the National Longitudinal Survey of Youth 1997 (NLSY97), I find strong evidence of layoff signaling early in a worker's career, with the magnitude of the signal effect gradually diminishing with pre-layoff experience. This suggests that, as a worker's pre-layoff experience grows and uncertainty about their ability decreases, employers gradually reduce the weight they place on the initial layoff signal. Consistent with previous literature (Pinkston 2003; Arcidiacono, Bayer, and Hizmo 2010), I find that effects vary by race and gender.

The rest of this paper is organized as follows. Section 1.2 provides a brief discussion of some of the related literature. Section 1.3 highlights the differences in wage dynamics around job loss for workers in the NLSY97 based on the type of job loss they experienced as well as based on their actual and relative ability levels. Section 1.4 presents the conceptual framework that motivates the empirical analysis of layoff signaling. Section 1.5 develops and empirically estimates a model of layoff signaling. Section 1.7 concludes.

[5]In a somewhat related study, Seim (2019) finds evidence that the size and persistence of earnings losses for laid-off workers do not appear to vary based on a worker's ability. The ability measure used in his study, however, is a worker's ability relative to the population, not relative to worker's with similar characteristics. When I repeat this empirical analysis using a worker's actual AFQT score, the results more closely match those found in the earlier study.

1.2 Background Information and Related Literature

This paper makes a number of contributions across various strands of literature. In this section, I highlight these contributions and discuss the how they tie in to the current state of the literature. I focus first on the methodological contributions to the different branches of the literature on understanding and identifying the ways in which employer learning affects wage dynamics over a worker's labor market career. I then discuss the contribution of the empirical findings to the literature on identifying the sources of the long-term earnings losses associated with involuntary job loss, as well as to the literature the effects early career job loss for young workers.

The empirical method developed in this paper expands on research that seeks to empirically identify asymmetric employer learning based on stigma effects of being laid-off. In a seminal paper, Gibbons and Katz (1991) - hereafter, GK - show that, under asymmetric information, distressed firms selectively lay off their lowest ability workers first, which sends a negative signal about those workers to prospective employers. GK attempt to identify these layoff signals by comparing outcomes across laid-off workers and workers who lost their jobs due to plant closure, where plant closures are assumed to be exogenous. While some researchers following GK's empirical approach find strong evidence of layoff signaling (Nakamura 2008; Kosovich 2010; Michaud 2018), others find little evidence in support of the theory (Grund 1999; Krashinsky 2002; Song 2007) or find evidence of layoff signaling only within specific populations (Gibbons and Katz 1991; Doiron 1995; Hu and Taber 2011). Stevens (1997) and Krashinsky (2002), however, provide evidence that differences in pre-job loss characteristics between workers who lost their jobs due to layoffs and plant closures, such as earnings trends and establishment size, may explain the effects found using GK's approach. Further, Lengermann and Vilhuber (2002) and Schwerdt (2011) show that relatively high ability workers leave distressed firms before those firms shut down. Because of this, workers who remain at the firm until it closes may be negatively selected. I avoid issues associated with using workers who lost their jobs due to plant closer as a comparison group for laid-off workers and instead compare how employers learn differentially about laid-off and non-laid-off workers.

I further expand upon the layoff signaling literature by allowing for layoff signals that change with experience and imperfectly reveal information about workers' abilities. That is, while GK's theoretical analysis, which is based on a two-period lemons model, is informative, it is limited in the sense that it does not take a stand on how layoff signals evolve with experience after a layoff, or how the interpretation of these signals change if some workers are laid-off for non-productivity related purposes, such as due to seniority rules.[6] Empirical studies on the signaling role of layoffs are generally agnostic about the evolution of signals with experience and the idea that layoff signals may be a form of statistical discrimination. A notable exception is Michaud (2018), who analyzes a model that allows for employers to inaccurately believe that some high-ability workers are low-ability due to a layoff signal and to correct their beliefs over time as they learn the worker's true type. Michaud (2018) empirically tests her model by comparing long-run effects of job loss for workers who lost their jobs due to layoff and plant closure, in an event study framework. I expand upon Michaud (2018) by analyzing the changing returns to a worker's relative ability with post-layoff experience, which allows for a more precise treatment of the learning process than an event study model. Additionally, unlike the current study, data limitations inherent to the PSID lead Michaud (2018) to treat layoffs and firings as indistinguishable events, which likely overstates the signaling effect of layoffs.[7]

This paper also contributes to the growing literature on empirically identifying employer learning that has grown out of the seminal symmetric learning models of H. Farber and Gibbons (1996) and Altonji and Pierret (2001) - hereafter FG and AP, respectively. The FG and AP models are based on the idea that employers form initial beliefs about each worker's ability based on a set of time-invariant, observable characteristics, such as education. As a worker's labor market experience increases, employers update their beliefs based on noisy output signals the worker sends each period. As a result, ability correlates, such as test scores, that are not observed by prospective employers should be uncorrelated with a worker's early career wages, conditional on observed characteristics. As the worker's labor market experience increases, however, ability correlates should become increasing correlated with the worker's wages due to employers learning the worker's true ability. AP's employer learning and statistical discrimination (EL-SD) model also shows that the more employers learn about a worker's ability, the less they rely on on observable characteristics in determining the worker's wage. FG and AP test their models using data from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). They examine returns to observable characteristics, as well as each worker's Armed Forces Qualifying Test (AFQT) score, a measure of their general aptitude

[6]GK do acknowledge that alternative layoff reasons, specifically seniority based layoff rules, likely affect their empirical analysis. They attempt to get around this issue by focusing on white-collar workers who are less likely to be laid-off due to seniority rules.

[7] The displaced worker literature generally excludes workers who are fired with cause from primary estimation samples due to endogeneity concerns regarding the type of worker who gets fired from a job (an activity that generally has higher fixed costs for firms (Oyer and Schaefer 2000)). See also Postel-Vinay and Turon (2013), Acharya, Baghai, and Subramanian (2013), Davis and Haltiwanger (2014), Haltiwanger, Scarpetta, and Schweiger (2014), and Mukoyama and Osotimehin (2019) for additional discussions on firing decisions/costs.

that it not observed by the employer, and find support for their theoretical predictions.[8] To the best of my knowledge, I am the first to merge aspects from the FG and AP models into a dynamic framework that allows for an imperfect layoff signal to be identified as a form of dynamic statistical discrimination that changes with both pre- and post-layoff experience.

In bridging the gap between the empirical employer learning literature and the layoff signaling literature, I also contribute to the work on empirically identifying asymmetric employer learning more broadly. While the theoretical foundation of asymmetric employer learning is well-established,[9] tractable empirical tests of asymmetric employer learning, outside of the layoff signaling literature, are a relatively new development and obtain mixed conclusions. Schönberg (2007a), for instance, develops a two-period theoretical model of asymmetric employer learning and derives predictions related to the return to ability with current tenure, relative to overall experience. She tests the predictions using an employer learning model based on AP and finds little evidence of asymmetric employer learning.[10] On the other hand, Pinkston (2009) develops an empirical asymmetric employer learning model in which asymmetric information is passed between employers in a worker's current employment spell, as opposed to being specific to a unique employer. He finds evidence that asymmetric learning has at least as large of an effect on wages as public learning during an employment spell.[11] While these earlier models allow for the identification of asymmetric employer learning in general, they do not specifically provide a means of addressing certain specific predictions from the literature, such as the signaling role of layoffs. I compliment earlier asymmetric employer learning studies by providing a general empirical framework for assessing predictions from the theory that previous models do not identify. Further, while the main empirical specification I study is based solely off of the FG and AP models, I show that it can be augmented to match the empirical specifications used in these earlier studies, which allows different predictions of the theory to be tested simultaneously.

[8]Other authors have expanded the general employer learning model in a number of ways. Lange (2007) modifies AP's EL-SD model to allow for the speed of employer learning to be structurally identified. Mansour (2012) expands on the AP model to test for differences in employer learning across initial occupation. Arcidiacono et al. (2010) and Light and McGee (2015a) break AP's NLSY79 sample into two separate samples based on highest education level attained (high school and college). Light and McGee (2015b) and Petre (2018) adjust these models to test the importance of different skill dimensions (ASVAB component test scores) and ability types (cognitive versus non-cognitive).

[9]See Waldman (2012) for a review

[10]Zhang (2007) extends Schönberg (2007a)'s theoretical model to three periods and finds evidence in support of asymmetric employer learning based on his model's predictions.

[11]Additional empirical tests for asymmetric employer learning include Devaro and Waldman (2012), L. B. Kahn (2013), Michaud (2018), Bates (2019), Fan and DeVaro (2020), and Cohn et al. (2021) all of which find evidence in support of asymmetric employer learning, but do so outside of the FG and AP framework. Additionally, while not specifically focused on asymmetric learning, the results of Mansour (2012) suggest that asymmetric employer learning exists at least between employers across occupations.

In addition to the methodological contributions, the empirical findings of this paper also make a number of contributions. More specifically, this paper contributes to the large body of work focused on understanding, and empirically identifying the causes of the long-term earnings losses associated with involuntary job loss, which have been well documented (see e.g. Jacobson et al. 1993; Couch and Placzek 2010). While numerous theories have been proposed over the past several decades,[12] a recent study by Lachowska et al. (2020) is the first to provide definitive empirical evidence on the mechanisms behind the long-term cost of job loss for long-tenure displaced workers (those with at least six years of pre-job loss tenure). The authors find that the earnings losses of long-tenure displaced workers are driven by wage losses that do not recover over time because (1) workers lose the specific human capital they had accumulated over their tenure with the firm and (2) they lose the benefits of a good quality job match that may have taken years to find. My work compliments Lachowska et al. (2020) by establishing empirical support for a mechanism behind the long-term costs of job loss for young workers who tend to be mechanically excluded from long-tenure displaced worker samples due to low levels of tenure driven by insufficient labor market experience and high levels of job-to-job mobility. Unlike long-tenure workers, younger workers are less likely to be affected by the loss of specific human capital (due to low levels of tenure) or the loss of a high quality job match (due to low levels of experience). As a result, I am able to directly assess both the overall contribution of layoff signaling to the long-run costs of involuntary job loss, and how this contribution changes with experience.

My empirical findings also contribute to the literature on the long-run effects of job loss for young workers more generally. Previous studies on the effects of job loss for young workers have generally found that earnings losses for young workers are dramatic in the period following job loss (though generally of smaller magnitude than for older workers), and tend to persist for at least five years following displacement, though these losses tend to taper out far faster than for older workers (Kletzer and Fairlie 2003; Barnette et al. 2021). Having a better understanding of the impact of events that occur early in a career have on young workers' long-term labor market outcomes is especially important given the vast literature on the long-run effects of early labor market conditions (Mroz and Savage 2006; von Wachter and Bender 2006).[13]

[12] Theories behind the long-term costs of job loss include lost industry/occupation specific capital (Topel 1990; Fallick 1993; Neal 1995); forgone human capital while unemployed (Burdett et al. 2020); skill mismatches upon reemployment (Nedelkoska, Neffke, and Wiederhold 2015; Kostol 2017); loss of position on career/occupation job ladders (Krolikowski 2017; Forsythe 2020); and costly post-job loss search (Jarosch 2021), to name a few. See Carrington and Fallick (2017) for a comprehensive review.

[13] Additional examples of studies examining the long-run effects of initial labor market conditions include L. B. Kahn (2010), Hershbein (2012), Altonji, Kahn, and Speer (2016), Liu, Salvanes, and Sørensen (2016), Schwandt and von Wachter (2019), and Arellano-Bover (2021). See also Rothstein (2020) who studies the effects of labor market entry after the Great Recession.

1.3 Wage Dynamics Around Job Loss

To motivate the main analysis of this paper, this section highlights differences in wage dynamics around job loss for different groups of young workers in the NLSY97, as well as inherent differences in the composition of these groups in terms of both observable and unobservable (by employers) characteristics. This is done for number of reasons. First, differences in wage dynamics around layoffs relative to plant closures are used to show that plant closures fail to provide a valid counterfactual when attempting to identify layoff signaling based on the difference in the effects of a layoff versus a plant closure. Second, differences, or lack there of, in wage dynamics around layoffs for workers with above- versus below-average ability (as proxied by AFQT score) are shown in order motivate the creation of a relative ability measure based on the measure created by H. Farber and Gibbons (1996). And third, differences in wage dynamics around layoffs for workers with above- versus below-average relative ability (as proxied by residual AFQT score) provide initial suggestive evidence that layoff signaling may plausibly play a role in determining the long-run cost of a layoff, which in turn provides motivation for the conceptual setting that will guide the main analysis of this paper.

Before proceeding further, it is worth reiterating that, as was mentioned in the introduction to this paper, if it is the case that employers are selectively laying off lower ability workers first, then event study models are likely unsuitable for causally studying the effects of job loss for young workers. That is, event study models are not identified in the presence of inherent differences in unobservable trends between workers who leave or are let go from firms relative to those who stay. This can lead to biases in the estimated effects of job loss (von Wachter and Bender 2006). Specifically, if some laid-off workers are negatively selected on the basis of ability, then the any estimated effects will be biased away from zero as the estimated effects pick up not only the wage dynamics associated with the true layoff effect, but also how the difference in the average return to ability between laid-off and non-laid-off workers changes over time due to employer learning. That is, even had the average laid-off worker not been laid-off, they would have had increasingly lower wages than the average non-laid-off worker over time due to employers learning that the ability of the laid-off worker is lower than it is for the average non-laid-off worker. As such, the effects estimated in this section should be considered as descriptive measures of general patterns of wage dynamics around job loss, and should be considered causal.

1.3.1 The NLSY97 Data

The data used in this study come from the 2017 release of the NLSY97, and is made up of workers who entered the labor market between January 1997 and December 2007.[14] The sample is first partitioned into two groups of workers, those who report having experienced an involuntary job loss at some point after they first enter labor market, and those who do not. As in Michaud (2018), I follow Gibbons and Katz (1991) in mapping involuntary separations into distinct categories based on whether the separation was due to plant closure or layoff. Additionally, workers who are terminated with cause are separated out from the layoff and plant closure samples, which is a distinct advantage to the NLSY97 data relative to the PSID data used by Michaud (2018) or DWS data used by GK.[15]

The analysis sample consists of a quarterly panel of 3,653 unique individuals. Of this sample, 753 individuals make up the layoff sample and 227 individuals make up the plant closure sample.[16] Table 1.1 compares baseline observable characteristics between laid-off workers and stably-employed workers (columns 1-2) and between laid-off workers and workers who lost their jobs due to plant closures (columns 4-5). For each reported variable, differences in mean values within each respective comparison group are reported in columns 3 and 6 of Table 1.1, along with the associated t statistics. The purpose of this table is highlight difference in baseline observable characteristics across different samples of workers that are likely associated with difference in unobservable determinants of productivity, specifically ability.

Workers in the layoff sample are significantly less likely to be female compared to both the non-job loss and the plant closure samples, and are more likely to be black or Hispanic compared to the the non-job loss sample. Additionally, workers in the layoff sample are 4-5 months older on average than workers in the plant closure sample but are more than nine months younger than workers in the non-job loss sample on average. This corresponds to baseline differences in years of completed schooling, with workers in the non-job loss sample having over a year more education on average than those in the layoff sample, who average only slightly more years of schooling than workers in the plant closure sample. Breaking down the differences in education further, approximately 63 percent of workers in the layoff sample have 12 or

[14] Workers who enter the labor market after the start of the Great Recession are excluded from this analysis to avoid confounding factors that may arise when studying job loss among workers scarred by entering the labor market during or immediately after the recession (L. B. Kahn 2010; Rothstein 2020).

[15] Full details on sample construction, statistics, and the construction of key variables variables can be found in Appendix A.1, while a formal description of the methods and criteria used to identify each of the job loss samples is described in Appendix A.1.1.

[16] An additional 518 individuals make up the fired worker sample. As the main focus of this study is on layoff signaling, the effects of being fired with cause are not presented but are available upon request. See the explanation in Footnote 7 for further references.

fewer years of education, which is similar to the plant closure sample, but far greater than the 41 percent of workers in the non-job loss sample 12 or fewer years of education. Conversely, while only 17 percent of workers in the layoff sample have 16 or more years of education, nearly 40 percent of workers in the non-job loss sample have at least 16 years of education, with 12 percent having strictly more than 16 years of education compared to only 5 percent of the layoff sample.

Table 1.1: Entry Quarter Summary Statistics By Sample Type

	Layoff and Stably Employed Samples			Layoff and Plant Closure Samples		
	Layoff Sample	Control Sample	Diff (t-stat)	Layoff Sample	Closed Sample	Diff (t-stat)
Worker Characteristics						
Female	0.36	0.55	-0.19***	0.36	0.47	-0.11**
	[0.48]	[0.50]	t =-9.14	[0.48]	[0.50]	t =-2.86
Black	0.31	0.21	0.11***	0.31	0.26	0.06+
	[0.46]	[0.41]	t =5.56	[0.46]	[0.44]	t =1.76
Hispanic	0.22	0.19	0.03+	0.22	0.26	-0.04
	[0.42]	[0.39]	t =1.65	[0.42]	[0.44]	t =-1.16
Asian	0.02	0.03	-0.01	0.02	0.02	0.00
	[0.14]	[0.16]	t =-1.31	[0.14]	[0.13]	t =0.10
Urban	0.80	0.78	0.02	0.80	0.80	-0.01
	[0.40]	[0.42]	t =0.98	[0.40]	[0.40]	t =-0.21
Age (*years*)	20.25	21.02	-0.77***	20.25	19.89	0.36*
	[2.18]	[2.25]	t =-8.27	[2.18]	[2.14]	t =2.18
Education (*years*)	12.79	13.89	-1.10***	12.79	12.67	0.11
	[1.89]	[2.21]	t =-13.11	[1.89]	[1.79]	t =0.82
Education (*groups*)						
<12 Years	0.12	0.06	0.055***	0.12	0.14	-0.021
	[0.32]	[0.24]	t =4.28	[0.32]	[0.35]	t =-0.83
12 Years	0.51	0.35	0.165***	0.51	0.47	0.047
	[0.50]	[0.48]	t =7.91	[0.50]	[0.50]	t =1.24
13−15 Years	0.19	0.19	0.001	0.19	0.23	-0.038
	[0.39]	[0.39]	t =0.09	[0.39]	[0.42]	t =-1.20
16 Years	0.12	0.27	-0.151***	0.12	0.15	-0.022
	[0.33]	[0.45]	t =-9.80	[0.33]	[0.35]	t =-0.83
>16 Years	0.05	0.12	-0.072***	0.05	0.02	0.034**
	[0.22]	[0.33]	t =-6.67	[0.22]	[0.13]	t =2.87
Employment Characteristics						
Wage	11.72	13.22	-1.50***	11.72	10.50	1.22***
	[7.16]	[8.97]	t =-4.62	[7.16]	[3.92]	t =3.32
Earnings/Quarter	4590.60	5342.20	-751.60***	4590.60	4020.50	570.11**
	[3674.30]	[4507.46]	t =-4.54	[3674.30]	[2318.19]	t =2.80
Hours/Quarter	389.41	403.50	-14.08*	389.41	384.44	4.97
	[151.08]	[159.06]	t =-2.17	[151.08]	[151.87]	t =0.43
Size of Employer	330.22	367.94	-37.73	330.22	172.42	157.80*
	[1465.39]	[1681.59]	t =-0.58	[1465.39]	[495.23]	t =2.52
Observations	753	2155		753	227	

Note: Standard deviations in brackets.
Source: Author's tabulations of NLSY97 data. See Section 1.3.1 for information regarding the construction of the sample presented here.

The difference in baseline characteristics between the job loss sample and the non-job loss sample are significant, and failure to account for these differences will likely complicate any attempt to identify layoff signaling by directly comparing outcomes across groups, especially if there are inherent differences in unobservable characteristics as well. That said, Table 1.1 is meant to specifically highlight mechanical differences between each group of workers that are likely to be accounted for by employers at labor market entry. Given this, it should not be surprising to see that workers in the layoff sample on average have lower earnings upon labor market entry compared to workers in the non-job loss sample, who tend to work more hours and at higher wages. What is perhaps more interesting is that, despite fairly similar baseline characteristics, workers in the layoff sample have significantly higher earnings at baseline on average compared to workers in the plant closure sample, with a majority of this difference due to having higher wages. One possible that this difference in wages is due to the fact that workers in the plant closure sample tend to work for significantly smaller employers on average, compared to workers in the layoff sample. This explanation seems plausible given evidence from a number of recent studies that the size of a worker's first employer is an important determinant of labor market career trajectories (e.g., Moscarini and Postel-Vinay 2012; Arellano-Bover 2020).

1.3.2 Event Study Wage Dynamics

I begin by illustrating the wage dynamics around job loss separately for workers who lose their jobs due to plant closure and those who lose their jobs due to layoff. This is done to establish a comparison point with studies that examine the signaling role of layoffs based on the empirical approach of Gibbons and Katz (1991). These studies rely on the assumption that plant closures are "exogenous" separations, while layoffs are "selective/endogenous" separations (Michaud 2018), and any differences in the overall effects between the two groups must be attributed to the endogenous selection in the layoff sample. Essentially, the assumption is that layoffs and plant closures should have approximately the same effect if both are exogenous job separations, but if layoffs send a negative information signal about a worker's ability, then the effects for layoffs should be worse than for plant closures.

I use the event study framework of Jacobson et al. (1993) to estimate the effect of job loss in period $t-k$ based on the following model:

$$Y_{it} = \alpha_i + \gamma_t + \mathbf{X}_{it}\beta_1 + \mathbf{Z}_{ie_0}\beta_2 + \sum_{k\geq -2}^{6} D_{it}^k \delta^k + \varepsilon_{it}, \tag{1.1}$$

where the outcome Y for individual i in period t depends on a worker specific fixed effect α_i; a time effect γ_t, specifically a vector of calendar year indicators; worker demographics $\mathbf{X}$ (including a quadratic in experience, the interaction of years of education and year indicators, and indicators for race, sex, part-time status, union membership, and urban residence, all interacted with experience and a year trend); characteristics of the worker's first employer when they entered the labor market $\mathbf{Z}_{e_0}$ (including the log of the employer's size and two-digit industry and occupation codes, all interacted with experience and a year trend); the effect of the k^{th} year relative to job loss δ_k, where D^k is an indicator for the k^{th} year relative to job loss; and a stochastic error term ε.[17] Identification in this type of model comes from the assumption that, absent job loss, individuals in the job loss group would have had similar outcomes to those in the control group. A simple test of this assumption comes from looking at the coefficients on the pre-job loss year indicators, in this case δ_{-2} and δ_{-1}. If no pre-job loss trends are found, these coefficients should both be near zero, indicating that wage trends for these workers were otherwise comparable to those of non-job losers prior to job loss.[18]

If employers believe that workers who are selectively separated from their previous employer are adversely selected but workers who are exogenously separated are not, then we should see more severe job loss effects for the selectively separated group relative to the plant closure group as the prior should be believed to be of lower ability on average than the latter. Figure 1.1 shows the log wage dynamics around layoff (1.1a) and plant closure (1.1b) for young workers. Relative to stably-employed workers, laid-off workers experience initial wage losses of around 10-13 percent, followed by a slow recovery to wage losses of five percent six years after being laid-off. While the post-job loss effects for workers in the plant closure sample appear to be as bad or worse than for those in the layoff sample, it is clear that the assumption of common trends is violated as these workers experience sizable wage reductions in the year(s) prior to job loss, a phenomenon also noted by Stevens (1997) who analyzes the effects of job loss using data from the PSID.[19]

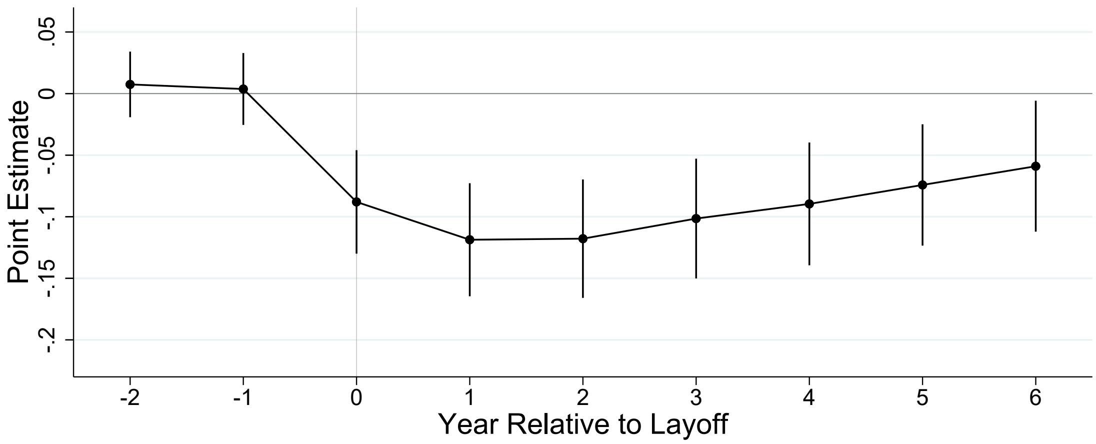
[17]Additional controls include indicators for the year an interview takes place, and indicators for job losses occurring prior to 2004 or after the start of 2008 which are used to capture any general trends in the costs of job loss associated with losing any job during those time periods, regardless of reason.

[18]Note that only workers with more than two years of pre-job loss experience are used when estimating this event study model, and that all periods prior to the two years before job loss have implicitly been set to have a zero estimated coefficient on a year relative to job loss indicator. Forcing a zero coefficient for at least one of the pre-job loss periods is required for identification in these types of models.

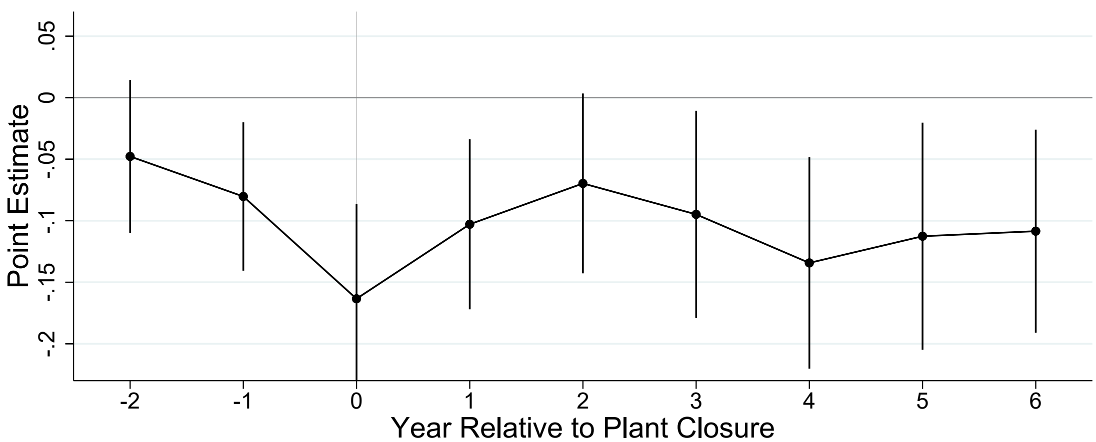
[19]Interestingly, Michaud (2018) uses a similar estimation strategy as presented here with data from the PSID, but does not document any pre-job loss trends for her sample of workers who lost their job due to a plant closure. It is possible that the differences between Michaud (2018) and Stevens (1997) in this regard are due in part to slightly different sample definitions, as well as different time frames studied in each of the samples.

Figure 1.1: Effects of Job Loss by Type

(a) Layoff Effect on Log Wage



(b) Plant Closed Effect on Log Wage



Notes: Sub-figures (a) and (b) report the estimated δ^ks - effects of job loss on log wages for workers who lose their job due to layoff or plant closure, respectively, pooled at the year level - based on Equation (1.1). Whiskers denote 95-percent confidence intervals based on standard errors clustered at the individual worker level. All workers in the job loss sample have an implied zero coefficient on an indicator for more than two years prior to their respective form of job loss (not shown). All workers with two years of experience or less at the time of job loss are excluded from these regressions to ensure identification.

Given that the common trends assumption is not violated in the pre-layoff period for laid-off workers, it cannot be reasonably assumed that workers who lose their jobs in a plant closure are otherwise comparable to laid-off workers prior to job loss, and thus cannot serve as a valid counterfactual after job loss. Thus, even if we were to assume that event study models of young worker job loss could be estimated without bias, the lack of a valid counterfactual would still prevent the signaling role of layoffs from being causally identified.

1.3.3 Post-Layoff Wage Dynamics By Ability

While using workers who lose their job due to plant closure proved to be an ineffective way to assess whether laid-off workers are adversely selected, a more useful approach may be to split the sample of laid-off workers by ability and compare the long-term layoff effects for each group, thus looking for difference in the long-run effects within the layoff sample instead of relative to a different form of job loss. To accomplish this, I exploit the unique information on pre-market skills in the NLSY97, specifically the age adjusted AFQT scores created by Altonji, Bharadwaj, and Lange (2012a). AFQT scores, which are derived from the math and reading portions of the ASVAB test, have been a standard measure of a worker's productive ability used in the literature dating back to Neal and Johnson (1996). As discussed by Lange (2007), AFQT scores provide researchers a plausible measure of a worker's productive ability that is not generally observable to employers, and thus capture a component of ability that employers must gradually learn about.

Table 1.2 provides descriptive statistics on the standardized AFQT scores for the non-job loss, layoff, and plant closure samples. Both job loss samples have average and 25th percentile scores that are significantly lower than the corresponding scores for the non-job loss sample. This should not come as a surprise given differences in the distribution of educational attainment across these groups shown previously in Table 1.1. What is perhaps more surprising that is while the 75th percentile score among the plant closure sample is lower than the 75th percentile score among the non-job loss sample (0.744 and 0.896, respectively), it is significantly larger than the 75th percentile score in the layoff sample, which is only 0.527. Given similar baseline characteristics between the plant closure and layoff sample, this difference could be an indication that some laid-off workers are being selectively laid-off due at least in part to their productive ability.

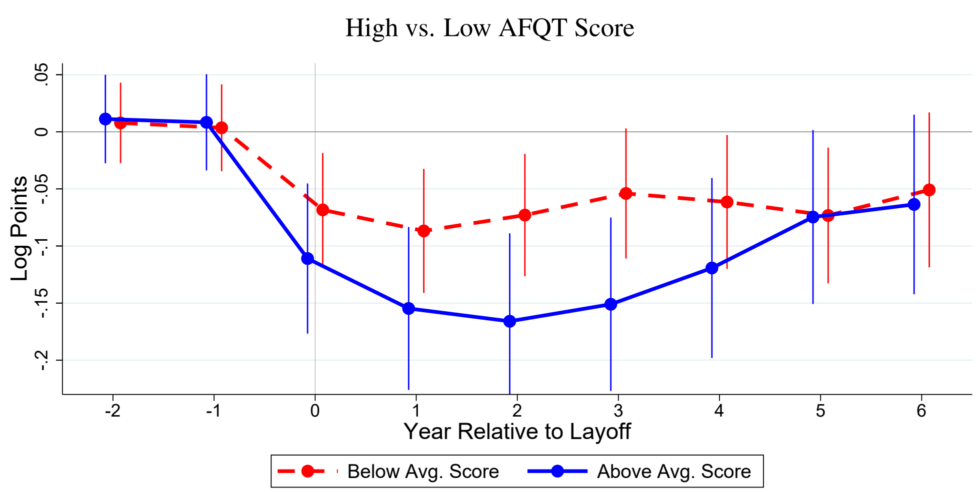
Table 1.2: Standardized AFQT Score Summary Stats By Sample Type

Sample	Mean	Std. Deviation	25^{th} Percentile	75^{th} Percentile
Non-Job Loss Sample	0.146	0.976	-0.491	0.896
Layoff Sample	-0.253	1.001	-1.037	0.527
Closed Sample	-0.118	1.038	-0.929	0.744
Total	0.031	1.001	-0.670	0.800

Source: Author's tabulations of NLSY97 data. See Section 1.3.1 for information regarding the construction of the sample presented here.

To see how differences in workers' ability levels contributes to differences in wages dynamics around the time of job loss, Equation 1.1 is re-estimated after first separating each of the job loss samples into groups by above- and below-average AFQT scores. Figure 1.2 displays the estimated δ^k's based on the regression described in Equation 1.1, broken down by whether an individual has above or below average ability based on their age adjusted AFQT score. The estimated post-layoff wage paths shown in this figure follow the general pattern found in Figure 1.1a and suggest that laid-off workers with above-average AFQT scores experience similar, if not worse, layoff effects as those with below-average scores. These post-layoff wage paths are remarkably similar to those found in Seim (2019)'s work looking at the effects of job loss by ability among Swedish workers.

Figure 1.2: Log Wage Effects of Layoff Grouped By Different Definitions of Ability Levels



Notes: Figure reports the estimated δ^ks - effects of a layoff on log wages, split by whether the worker is above or below average AFQT score, pooled at the year level - based on splitting Equation (1.1) by above or below average AFQT. Whiskers denote 95-percent confidence intervals based on standard errors clustered at the individual worker level. All workers in the job loss sample have an implied zero coefficient on an indicator for more than two years prior to their respective form of job loss (not shown). All workers with two years of experience or less at the time of job loss are excluded from these regressions to ensure identification.

While this may seem to suggest that adverse selection and signaling do not play a major role in determining the long-term effects of layoffs, it is important to consider that ability (AFQT score) is highly correlated with other characteristics that have been shown to greatly impact the long-term costs of job loss, such as education (H. Farber 2017). As such, it is possible that differences in the above- and below- average AFQT score groups that are unrelated to what employers are trying to learn about a worker are counteracting any signaling effect that may be present when the sample is separated in this way. Splitting the layoff sample based on a measures of a workers relative ability should mitigate any problem with confounding factors as this measure is orthogonal to other observable characteristics by construction.

In order to construct a measure of a worker's relative ability, I follow H. Farber and Gibbons (1996) in creating residual AFQT scores that are orthogonal to the characteristics observed by prospective employers when a worker first enters the labor market. To accomplish this, I first create a sample consisting of each individual's first period in the labor market (i.e. the first period they are employed and report a non-zero wage). Then, following FG, I define residual ability z_i^* as,

$$z_i^* = z_i - E^*(z_i \mid X_{i0}, \omega_{i0}),$$

where E^* is the linear projection of an ability measure z on a vector of observable characteristics X_{i0} and the worker's first period wage ω_{i0}. FG show that z^* is equivalent to observing employers' expectation error in a worker's ability for experience levels $t > 0$, and can be used by researchers to assess the effects of employer learning.[20] In practice, I regress each worker's AFQT score on a vector of observable characteristics and their first period wage, and then use the fitted values from this regression to calculate each worker's residual AFQT score.

The vector of observable characteristics used to create the residual AFQT scores contains five education dummy variables (< 12 years, 12 years, 13-15 years, 16 years, and 17+ years), an indicator for part-time status, the interaction of part-time status and each education dummy, indicators for race, sex, marital status, marital status interacted with sex, age in years (<18, 18-19, 20-21, 22-23, 23-24, 25+), birth year, current year and quarter, and the log of the worker's wage.[21] Additionally, I include indicators for the number of employees at each worker's employer (≤10, 11-50, 51-100, and 100+), as well as the interaction of each of these indicators with each worker's log wage. This is done to account for potential differences in an employer's ability to judge a worker's true ability based on the employer's size.

[20] Light and McGee (2015b) use a slightly different approach than FG. They regress their z measures only on the observable characteristics used in their model, leaving out the entry period wage. The advantage of that approach is that it does not require the entry period wage to be dropped in log-wage regression models, while still purging their ability measure of any correlation to observable characteristics. However, that approach does not purge the correlation between characteristics that are only observed by the employer and the ability measure, which will complicate the interpretation of the estimated return to ability over experience, since certain aspects of learning will be correlated with these observable characteristics, meaning their $\hat{z}$ is a biased measure of employers' true expectation errors. This issue is especially problematic in this context as the characteristics observable to employers are likely to have some form of correlation with the any job loss signal, which would make it impossible to distinguish between the signaling effect of the job loss and this correlation.

[21] Each worker's first period wage is included to serve as a proxy for any additional observable characteristics that are observed by the employer but are not found in the data. See H. Farber and Gibbons (1996) for further discussion. Residual AFQT scores that are constructed without the first period wage will be used for robustness checks of the primary specifications.

This regression accounts for roughly 40 percent of the variation in AFQT scores, which is noticeably lower than the R^2 value found in FG (53 percent of variation accounted for using their NLSY79 sample and nearly identical controls). This seems to be in line with recent empirical evidence from Altonji et al. (2012a), who find that the ability distribution has widened over time, and that demographic characteristics, such as race and gender, appear to play a less predictive role in an individual's AFQT score among the more recent cohort.[22] Lastly, while $AFQT_i^*$ is mean zero by construction, to make the measure comparable to the standardized AFQT scores, it is normalized to have unit-variance.

Table 1.3 provides summary statistics related to these residual AFQT scores, as well as the associated predicted AFQT scores for comparison, among workers in the non-job loss, layoff, and plant closure samples. While the mean residual AFQT score for each sample is closer to the sample mean than the mean standardized AFQT scores were, the average for the layoff sample is still significantly lower than that of the other two samples, both of which are slightly higher than, though statistically equal to the sample average. This stands in sharp contrast to the predicted AFQT scores reported in the lower half of Table 1.3 in which the distribution of scores for the layoff and plant closure samples appear nearly identical, and skewed far to the left of the distribution of predicted scores for the non-job loss sample. Taken together, this suggests that while observable characteristics account for the higher (lower) average AFQT scores reported by the non-job loss (plant closure) sample, the lower scores reported for the layoff sample cannot be accounted for based solely on observable characteristics. That is, on average, workers in the layoff sample have a lower overall ability than their observable characteristics would suggest. While this is merely descriptive, the fact that laid-off workers appear to be negatively selected on unobserved ability provides some justification for employers' to use an observed layoff as a signal of a worker's ability, resulting in layoffs acting as a form of statistical discrimination.

To see how differences in workers' ability levels contributes to differences in wages dynamics around the time of job loss, Equation 1.1 is re-estimated after first separating each of the job loss samples into groups by above- and below-average residual AFQT scores. Figure 1.3 displays the estimated δ^k's based on the regression described in Equation (1.1), broken down by whether an individual has above or below average relative ability based on their residual AFQT score (1.3b) and whether they have above or below

[22] It is also possible that this smaller relationship between observable characteristics and AFQT scores could be related to the recent evidence that the return to cognitive ability has generally been decreasing over the past few decades (e.g., Castex and Dechter 2014; Beaudry, Green, and Sand 2016), though evidence from other studies using more updated data, such as Ashworth et al. (2020), among others, suggest that this may not be the case.

Table 1.3: Residual & Predicted AFQT Score Summary Stats By Sample Type

Residual AFQT Score				
Sample	Mean	Std. Deviation	25th Percentile	75th Percentile
Non-Job Loss Sample	0.008	0.960	-0.591	0.661
Layoff Sample	-0.085	1.021	-0.722	0.665
Closed Sample	0.053	1.118	-0.697	0.865
Total	-0.011	0.988	-0.642	0.675

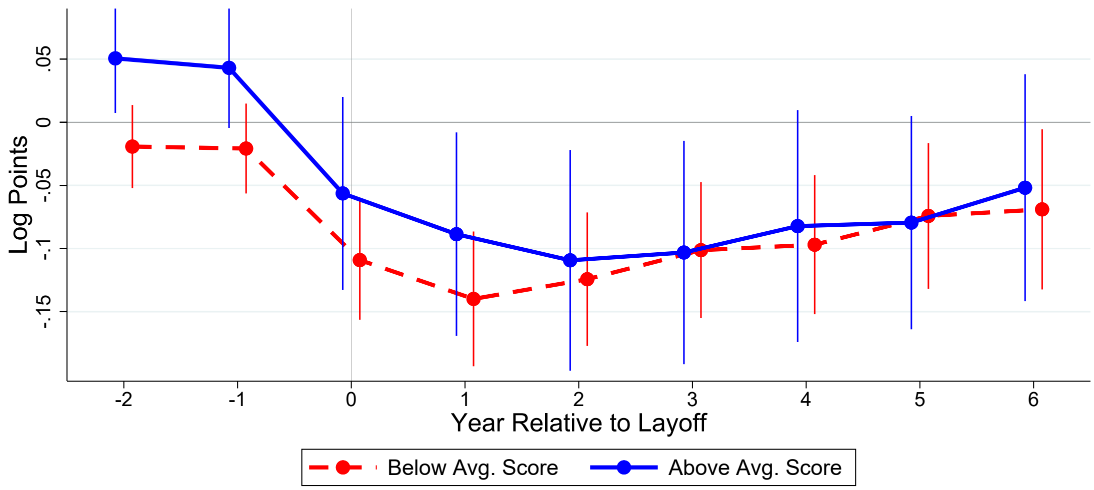
Predicted AFQT Score				
Sample	Mean	Std. Deviation	25th Percentile	75th Percentile
Non-Job Loss Sample	0.219	1.007	-0.477	1.116
Layoff Sample	-0.293	0.948	-0.979	0.318
Closed Sample	-0.249	0.920	-0.906	0.377
Total	0.062	1.014	-0.669	0.863

Source: Author's tabulations of NLSY97 data. See Section 1.3.1 for information regarding the construction of the sample presented here.

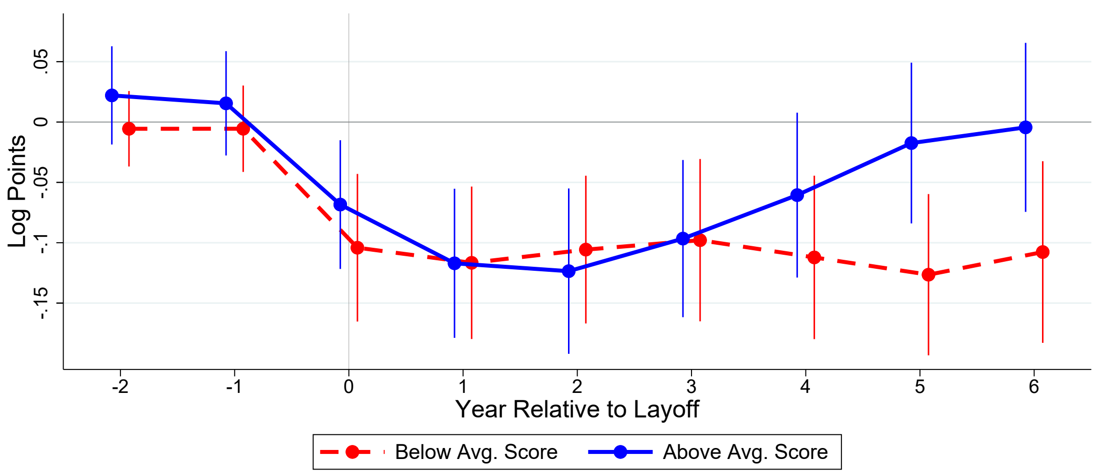
average predicted ability based on their predicted AFQT score (1.3a). In terms of predicted AFQT score, not only due the post-layoff effects appear identical between the two groups of workers, but the workers in the above average predicted AFQT score group appear to experience significant positive pre-layoff wage effects. One possible explanation for these positive pre-layoff effects could be that these are workers who employers incorrectly assumed were high ability workers, hence the higher wages, but were let go after being revealed to negative residual ability. The results from Figure 1.3b offer some potential support for this hypothesis. While workers with above- and below-average residual AFQT scores experience nearly identical initial wage losses following a layoff, workers with above-average residual AFQT scores gradually recover to the point where the effects of the layoff are statistically indistinguishable from zero after five years, while workers with below-average residual AFQT scores experience no noticeable recovery over the same period. Given that there appear to be no pre-layoff effects when splitting the sample along residual AFQT score, it is reasonable to infer that the pre-layoff effects found for the above-average predicted AFQT portion of Figure 1.3a are driven by workers with below-average residual AFQT scores, since only the mechanism partitioning the sample changed between estimations.

Figure 1.3: Log Wage Effects of Layoff Grouped By Residual and Predicted of Ability Levels

(a) High vs. Low Predicted AFQT Score



(b) High vs. Low Residual AFQT Score



Notes: Sub-figures (a) and (b) report the estimated δ^ks - effects of a layoff on log wages, split by whether the worker has an above or below average residual AFQT score or predicted AFQT score, respectively, pooled at the year level - based on splitting Equation (1.1) by above or below average ability definitions. Whiskers denote 95-percent confidence intervals based on standard errors clustered at the individual worker level. All workers in the job loss sample have an implied zero coefficient on an indicator for more than two years prior to their respective form of job loss (not shown). All workers with two years of experience or less at the time of job loss are excluded from these regressions to ensure identification.

Returning focus to Figure 1.3b, while there are likely numerous possible explanations for the divergent post-layoff wage paths by residual AFQT score, the presence of these divergent paths is consistent with a layoff signaling story wherein employers statistically discriminate among workers on the basis of an observed layoff. That is, if employers are unable to distinguish between above- and below-average relative ability workers at the time of layoff, they assign all laid-off workers the same amount of negative information, based on the average expected ability of laid-off workers in the population. This results in similar wage losses immediately following a layoff for all workers independent of relative ability. Then, over time, as employers learn each worker's true ability, wages for the above-average relative ability workers recover as employers correct their inaccurate beliefs from the layoff signal, while wages remain "suppressed" for below-average relative ability workers as employers confirm their belief based on the signal. Despite this, the wage dynamics in Figure 1.3b only confirm that workers with above-average residual ability do a better job at recovering from the effects of a layoff than do workers with below-average residual ability. There are likely a multitude of theoretical justifications for why workers with above-average residual ability recover from layoffs better than those who are below-average, such as better potential to rise quickly through job ladders. Thus, in order to parse out the extent to which these divergent recovery paths are indeed driven by some form of layoff signaling, it is necessary to develop specific predictions that distinguish layoff signaling from other possible explanations.

Before continuing further, however, it is important to highlight one additional consideration. Specifically, if we are to believe that prospective employers are correcting inaccurate beliefs about a worker over time after the worker is laid-off, then we should also have reason to believe that prospective employers are learning about the worker prior to the layoff. If this is the case, the extent to which laid-off workers appear to be negatively selected should decrease with pre-layoff experience since the extent to which a current employer holds an information advantage over prospective employers should decrease as prospective employers gain additional information. Table 1.4, which breaks down a number of key worker characteristics by years of pre-layoff experience, supports this general idea. While there does not appear to be a significant relationship between years of pre-layoff experience and AFQT or predicted AFQT scores, the same does not appear to be the case for residual AFQT score, which appears to increase monotonically with pre-layoff experience. Specifically, average residual AFQT scores increase gradually from -0.296 for workers with one year of pre-layoff experience to .086 for workers with 10+ years of pre-layoff experience. While purely descriptive, this suggests that it is plausible that pre-layoff learning could play a role in gradually reducing the amount of information conveyed by a layoff. This pre-layoff experience-learning dynamic is going to play an important role in how layoff signaling is identified by the method developed in this study.

In the next section, I lay out a conceptual framework of layoff signaling that yields a number of predictions that are associated specifically with layoff signaling. Then, in Section 1.5 , I develop an empirical

Table 1.4: Summary Statistics For Layoff Sample By Years of Pre-Layoff Experience

	1	2 – 3	4 – 5	6 – 7	8 – 9	10+	Total
Pre-Layoff Wage	**12.4**	**14.2**	**16.1**	**15.9**	**17.8**	**18.4**	**15.5**
	[4.94]	[6.21]	[10.76]	[6.18]	[7.61]	[9.93]	[7.98]
Pre-Layoff Hours/Week	**40**	**39.6**	**40.4**	**40.1**	**41.3**	**42.2**	**40.4**
	[6.73]	[7.09]	[7.15]	[7.96]	[7.47]	[8.56]	[7.42]
Female	**.315**	**.345**	**.403**	**.345**	**.357**	**.443**	**.364**
	[0.47]	[0.48]	[0.49]	[0.48]	[0.48]	[0.50]	[0.48]
Black	**.278**	**.354**	**.295**	**.319**	**.337**	**.266**	**.315**
	[0.45]	[0.48]	[0.46]	[0.47]	[0.48]	[0.44]	[0.46]
Hispanic	**.231**	**.194**	**.275**	**.159**	**.204**	**.291**	**.222**
	[0.42]	[0.40]	[0.45]	[0.37]	[0.41]	[0.46]	[0.42]
Tenure at Job Loss	**.766**	**1.33**	**2.06**	**2.33**	**3.01**	**4**	**2.04**
	[0.46]	[0.93]	[1.60]	[2.09]	[2.74]	[3.60]	[2.15]
Job Spell Length at Job Loss	**1.04**	**1.71**	**2.79**	**3.07**	**3.42**	**4.72**	**2.57**
	[0.48]	[1.04]	[1.91]	[2.61]	[3.30]	[4.03]	[2.51]
Standardized AFQT Score	**-.468**	**-.307**	**-.243**	**-.116**	**-.155**	**-.155**	**-.253**
	[1.00]	[0.96]	[1.00]	[1.07]	[0.97]	[1.00]	[1.00]
Above Avg. AFQT Score	**.343**	**.442**	**.477**	**.504**	**.5**	**.532**	**.461**
	[0.48]	[0.50]	[0.50]	[0.50]	[0.50]	[0.50]	[0.50]
Std. Predicted AFQT Score	**-.376**	**-.322**	**-.254**	**-.152**	**-.32**	**-.345**	**-.293**
	[0.97]	[0.99]	[0.92]	[0.90]	[0.97]	[0.91]	[0.95]
Above Avg. Predicted AFQT	**.324**	**.359**	**.396**	**.416**	**.347**	**.354**	**.368**
	[0.47]	[0.48]	[0.49]	[0.50]	[0.48]	[0.48]	[0.48]
Std. Residual AFQT Score	**-.296**	**-.131**	**-.105**	**-.0245**	**.0649**	**.086**	**-.0853**
	[1.05]	[0.94]	[1.02]	[1.09]	[1.06]	[0.99]	[1.02]
Above Avg. Residual AFQT	**.407**	**.495**	**.483**	**.549**	**.541**	**.582**	**.503**
	[0.49]	[0.50]	[0.50]	[0.50]	[0.50]	[0.50]	[0.50]
Average Year	**2003.7**	**2005.3**	**2007.8**	**2009.4**	**2011.3**	**2013.4**	**2007.5**
	[2.48]	[2.78]	[2.50]	[2.16]	[2.42]	[1.98]	[3.95]
Observations	108	206	149	113	98	79	753

Note: Standard deviations in brackets.
Source: Author's tabulations of NLSY97 data. See Section 1.3.1 for information regarding the construction of the sample presented here.

approach to test these predictions that exploits the differences in the composition of ability between laid-off and non-laid-off workers. This empirical approach allows layoff signaling to be identified without relying on the assumptions that are violated in the traditional event study design.

1.4 Implications of Layoff Signaling

In this section, I illustrate the implications for wage dynamics as a result of layoff signaling under a number of key assumptions that are based on established empirical results from different strands of the employer learning literature. This yields a number of testable predictions that relate GK style layoff signaling to a dynamic form of statistical discrimination.

1.4.1 Conceptual Setting

I rely on three key assumptions to establish the intuition behind a framework from which it is possible to conceptualize the wage dynamics associated with layoff signaling.

1. Employers are initially uncertain about a worker's true ability at labor market entry, but gradually learn as the worker gains experience.
2. A worker's current employer has weakly more information about the worker's ability than prospective employers, and has strictly more information early in a worker's career.
3. Both private and public information converge to full information as a worker's experience in the labor market increases.

The first assumption sets a baseline environment in which the extent to which employers rely on any single source of information to form their beliefs about a worker's ability is decreasing in the worker's labor market experience. The second assumption establishes the presence of asymmetric information about worker ability for at least some workers at any given time, and is necessary for a GK-type layoff signaling "game" to take place. The third, and strongest assumption, ensures that workers' wages gradually converge on their full information wage as their experience increases.[23,24] Given these baseline assumptions, a worker's ability at any level of labor market experience can be decomposed into a "known/predicted" portion that is based on the amount of public information available at the time, and an "unknown/residual" portion about which prospective employers must learn over time.

Since current employers always have at least weakly more information than prospective employers, the latter may rely on observed actions of the former, specifically the decision to selectively lay off a worker, to gain additional information about the worker's ability. The informativeness of any additional information from an observed layoff depends on the extent of the information advantage held by the current employer. Thus, after observing that a worker has been laid-off, public information about the worker's ability is updated to account for this new "layoff signal," which follows from a publicly known layoff rule. The layoff rule is assumed to be based on a worker's ability, as well as other possible components of the layoff decision, such as match quality, and thus serves as an imprecise signal. The weight that prospective employers place

[23] A weaker assumption would be to assume that the degree of information asymmetry between current and prospective employers is weakly decreasing with each year of labor market experience, and strictly decreasing for at least some. While this assumption would still lead to the predictions discussed later in this section, the stronger assumption provides a clearer illustration of the signaling effects and is maintained for simplicity.

[24] Prior studies, reviewed in Section 1.2, have found empirical support for the first and second assumption. While not explicitly stated, these studies generally rely on a form of assumption three as a consequence of the structure of public information, which is generally modeled as a continually improving process. See e.g. Pinkston (2009) for an example where the precision of the public signal is strictly increasing with experience.

on the layoff signal is based on the degree of information asymmetry about the worker's ability at the time of the layoff. Because prospective employers are unable to distinguish between workers who are laid-off due to low productive ability and those laid-off for non-ability related reasons, the layoff signal attributes to each worker the average amount of negative information expected to be conveyed based on the distribution of ability among laid-off workers.

The idea behind this setting is that, as employers gradually learn about a worker's ability, the changing nature of asymmetric information across employers, conditional on the amount of public information available, leads to repeated GK-type layoff signaling "games" at each level of a worker's experience. That is, at the end of each period a worker is in the labor market, any new public information available about their ability results in an updated public expectation of the worker's ability, which results in a new public expectation error for the worker. This new expectation error can be thought of as a worker's updated relative ability, which is simply the new difference between the worker's actual ability and the average ability of his or her peers with the same observable characteristics and amount of public information. Then, at the start of each new period the worker is in the labor market, employers without access to the private information are unable to distinguish between workers based on their updated relative ability levels. Thus, in each period, the underlying mechanism driving layoff signaling essentially resets based on any new public information gained the prior period, and is analogous to the general mechanism described in GK, except that, in this setting, the information asymmetry is in regard to a worker's updated relative ability, rather than the worker's initial relative ability at labor market entry.

Finally, because employers assign all laid-off workers the average amount of negative information contained in the layoff signal, even those for whom the layoff was the result of a large productivity shock, the signal is effectively a form of statistical discrimination, about which employers learn the accuracy of over time as more information becomes available. Further, since public information about a worker's ability would have converged to full information over time in the absence of a layoff, the longer a worker is in the labor market post-layoff, the more employers reduce the weight placed on the negative information conveyed by a layoff signal in favor of the additional public information revealed about each workers' ability since the layoff. Ultimately, this setting describes statistical discrimination in the form of layoff signals that evolves based on both pre- and post-layoff employer learning.

1.4.2 Empirical Predictions

The wage dynamics implied by this conceptual setting yield a number of predictions that I will take to the data. These predictions are driven by the size of the disparity between public and private information about a worker's ability and how this disparity evolves as a worker's experience increases. If this disparity is large, then prospective employers are expected to believe that a layoff contains informative negative information about a worker's ability. Conversely, if this disparity is small, then prospective employers are expected to believe that a layoff conveys little information about the worker's ability, and thus beliefs should not change based on the observed layoff. The empirical predictions based on this framework are summarized as follows.

(i) **A negative layoff signal initially disproportionately affects high ability workers.** Since the probability of being laid-off is never zero, some portion of the laid-off population will be high ability workers. When prospective employers assign high ability workers a negative layoff signal, their beliefs about these workers' ability levels move *further* from the truth than for non-laid-off high ability workers. Conversely, a negative layoff signal moves employers' beliefs *closer* to the truth for low ability workers relative to their non-laid-off peers.

(ii) **Following a negative layoff signal, employers update their beliefs about high ability workers faster than for similar non-laid-off workers.** When employers use output signals to update their beliefs about the ability of a non-laid-off worker, they weight each signal based on its underlying precision, which is increasing in experience. For workers for whom the layoff was an inaccurate signal, as employers increase the weight placed on output signals, they decrease the weight placed on the negative layoff signal. This combined effect results in a faster rate of change in employers' beliefs for high ability laid-off workers relative to that for similar non-laid-off workers.

(iii) **The magnitude of the initial layoff signal effect decreases with pre-layoff experience, regardless of ability level.** As a worker's labor market experience increases, the public signal becomes an increasingly precise signal of the worker's ability. As a result, the relative information advantage maintained by the worker's current employer decreases in experience. As this information advantage decreases, layoffs become an increasingly noisy signal of a worker's ability. As the noise in the layoff signal increases, prospective employers decrease the amount of weight they place on observing the signal, which decreases the initial layoff signal effect.

(iv) **The rate at which employers update their beliefs about high ability, laid-off workers decreases with pre-layoff experience.** When the initial effect of the layoff signal decreases, the extent to which employers must correct their beliefs as they observe additional output signals also decreases. The more pre-layoff experience a worker has, the less additional employer learning needs to take place. As such, the rate of learning about a laid-off worker's ability converges to that of non-laid-off workers as pre-layoff experience increases.

1.5 Empirically Identifying Layoff Signaling

In this section, I develop an empirical strategy that allows layoff signaling to be identified based on the empirical employer learning models of H. Farber and Gibbons (1996) and Altonji and Pierret (2001). AP's model provides a useful framework for assessing the presence of statistical discrimination by employers regarding a worker's pre-labor market ability (as proxied by their AFQT score). Unlike the standard event study design discussed in Section 1.3, the empirical approach developed in this section is designed to leverage the differences in the distribution of ability between the laid-off and non-laid-off worker samples for identification, thus directly accounting for the employer learning related bias that is present in the estimated effects from the event study design. Additionally, while the empirical approach pursued in this section extends the implications of AP's EL-SD model to account for layoff signaling, a formal extension of AP's model that allows for wage dynamics that depend on dynamic post-labor market entry statistical discrimination is pursued in VanderBerg (2021a) and includes specific derivations that formalize the empirical approach highlighted in this section.

1.5.1 Empirical Strategy

AP's empirical approach attempts to model the role of statistical discrimination in an employer learning process based on the relationship between each worker's unobserved productivity (α) and characteristics that are observable to employers at labor market entry (s,q).[25] Under the assumption that the expectation of α given s and q is linear in s and q,

$$\alpha = E(\alpha|s,q) + \tilde{\alpha} = \phi_s s + \phi_q q + \tilde{\alpha},$$

where $\tilde{\alpha}$ is the remaining error in employers' initial beliefs about the worker's productivity, and is analogous to the relative ability terms used in this study. This decomposition leads to the log wage process for workers with x years of experience given by

$$w_x = (\phi_s + \gamma)s + (\phi_q + \kappa)q + H^*(x) + E(\tilde{\alpha}|F_x) + \xi_x, \tag{1.2}$$

where $E(\tilde{\alpha}|F_x)$ represents the extent to which employers have learned about their initial expectation error based on the information available at x, denoted F_x; $H^*(x)$ represents the general wage return to experience and others factors that evolve with experience that are outside of the model; and ξ_x is idiosyncratic error.

[25] q variables are observed by employers but not the researcher, while s variables are observed by both the employer and the researcher.

To establish their main empirical implications, AP derive the coefficients from a regression of w_x on an s variable(s) and an ability proxy, say z, for workers with x years of experience, based on Equation (1.2):

$$E(w_x|s,z,x) = b_{sx}s + b_{zx}z + H^*(x), \tag{1.3}$$

where s, z, and q are reinterpreted as the components of s, z, and q that are orthogonal to $H^*(x)$.[26] Then, based on the omitted variables bias formula for OLS, AP show that the coefficients b_{sx} and b_{zx} are

$$\begin{aligned} b_{sx} &= (\phi_s + \gamma) + \Phi_{qs} + \Phi_{sx} \\ b_{zx} &= \Phi_{qz} + \Phi_{zx}, \end{aligned} \tag{1.4}$$

where Φ_{qs} and Φ_{qz} are the coefficients from a regression of $(\phi_q + \kappa)q$ on s and z, and Φ_{sx} and Φ_{zx} are the coefficients from a regression of $E(\tilde{\alpha}|F_x)$ on s and z. The components of b_{sx} and b_{zx} that vary with experience (and thus pick up employer learning), can be expressed as

$$\Phi_{sx} = \theta_x \Phi_s \text{ and } \Phi_{zx} = \theta_x \Phi_z,$$

where Φ_s and Φ_z are the coefficients from the regression of $\tilde{\alpha}$ on s and z, while $\theta_x \in [0,1]$ describes the extent to which employers have learned about $\tilde{\alpha}$, and is assumed to be non-decreasing in experience (see AP Proposition 1). Further, AP show that Φ_s can be expressed as $-\Phi_z \Phi_{zs}$, where Φ_{zs} is the coefficient from a regression of z on s and highlights the relationship between the ability correlate z and the extent to which employers use s to infer a worker's ability at labor market entry. Plugging this relationship into Equation 1.4, AP's Proposition 2 shows that,

$$\frac{\partial b_{sx}}{\partial x} = -\Phi_{sz} \frac{\partial b_{zx}}{\partial x}, \tag{1.5}$$

which describes the manner in which employer learning about z spills over onto the coefficients on variables used by employers to statistically discriminate among workers at labor market entry.

Equation 1.5 forms the basis for the empirical approach AP use to pick up employer learning and statistical discrimination. To empirically test the predictions of their model, AP estimate a log wage equation of the form

$$w_{it} = \mu_0 + \tau_t + \gamma_s s_i + \gamma_{sx}(s_i \times x_{it}) + \beta_z z_i + \beta_{zx}(z_i \times x_{it}) + f(x_{it}) + \beta_\Psi \Psi_i + \varepsilon_{it}. \tag{1.6}$$

For worker i in period t, log wages w_{it} depend on observable characteristics s_i (such as schooling), an ability measure z_i (AFQT score), experience x_{it}, time effects τ_t, controls Ψ_i, and an idiosyncratic error term ε_{it}.

[26] See AP's discussion following their Equation (4) for more details on this distinction.

The intuition behind using this regressing to assess AP's prediction is that, if employers rely on observable characteristics to infer a worker's ability at labor market entry, we should find that the estimated effect of AFQT score on log wages is small at low levels of experience ($\hat{\beta}_z \approx 0$), but increases as experience grows ($\hat{\beta}_{zx} > 0$). Conversely, for observable characteristics that are positively correlated with AFQT score, such as education, we should find that the estimated effect of these characteristics on log wages is initially large ($\hat{\gamma}_s > 0$), but gradually decreases with experience ($\hat{\gamma}_{sx} < 0$). Essentially, the interaction of x with the s and z variables is meant to capture the wage dynamics that result from Equation 1.5, and thus provide evidence that employers not only learn about a worker's true productivity as experience increases, but that they used easily observed characteristics to initially statistically discriminate against the worker at labor market entry.

The usefulness of the intuition behind AP's results for assessing the presence of layoff signaling comes from the fact that, as discussed in Section 1.4, if employers rely on an observed layoff as a negative signal of a worker's unobserved ability and update their beliefs accordingly, they are statistically discriminating against the worker based on this observation. While AP's EL-SD model was designed to address statistical discrimination based on characteristics fixed at labor market entry, relative to a worker's overall ability, the intuition behind the model can be adapted to account for statistical discrimination that occurs at some point after a worker has entered the labor market. This can be done by utilizing the initial setup proposed by H. Farber and Gibbons (1996), who modeled employer learning regarding the portion of a worker's ability that is uncorrelated with observable characteristics at labor market entry, $\tilde{z}_i$ (as proxied by residual AFQT scores), and allowing for changing post-labor market entry characteristics to affect the learning process related to this relative ability measure.[27]

In essence, unlike AP's static EL-SD model, the timing of the statistical discrimination based on a layoff signal is evolving over time, with each new period in which a layoff occurs acting as a unique application of the EL-SD model, conditional on the relevant pre-layoff learning. To use this intuition to test the predictions discussed in Section 1.4, the empirical log wage estimation model of AP shown in Equation 1.6 is modified to include a number of additional variables.[28] Specifically, to identify the changing nature of layoff signals,

[27] See VanderBerg (2021a) for a formal treatment of the wage dynamics associated with such an extension to AP's model.

[28] A formal approach to modeling the dynamics of layoff signaling would be to illustrate the bias in the OLS estimates from AP's regression equation (Equation 1.3) when the true wage processes varies dynamically with experience around the time of a layoff. This approach is pursued in VanderBerg (2021a), and considers the coefficients from the following regression of w_x on s and z^* for a worker with x years of experience,

$$E(w_x|s,z^*,x) = b_{sx}s + b_{zx}z^* + H^*(x).$$

this study estimates the following log wage model,

$$
\begin{aligned}
\omega_{it} = \; & \mu_0 + \tau_t + \gamma_s s_i + \gamma_{sx}(s_i \times x_{it}) + \beta_a AFQT_i^* + \beta_{ax}(AFQT_i^* \times x_{it}) + f(x_{it}) + \beta_\Psi \Psi_i \\
& + \beta_a^0 (D_{it} \times AFQT_i^*) + \beta_a^x (D_{it} \times AFQT_i^* \times Pre_i) + \beta_{ax'}^0 (D_{it} \times AFQT_i^* \times Post_{it}) \\
& + \beta_{ax'}^x (D_{it} \times AFQT_i^* \times Post_{it} \times Pre_{it}) + D_{it}\delta + f_D(Pre_i, Post_{it}) + \varepsilon_{it},
\end{aligned} \tag{1.7}
$$

which adds the interaction of residual AFQT score ($AFQT^*$) and an indicator for experiencing a layoff (D_{it}), the interaction of $AFQT^*$ and pre-layoff experience (Pre_i), the interaction of $AFQT^*$ and post-layoff experience ($post_{it}$), and the interaction of $AFQT^*$ and pre- and post-layoff experience. While it is the post-layoff experience profile that picks up the difference in employer learning following a layoff, it is necessary to account for learning that occurs with pre-layoff experience, and how this pre-layoff learning results in a smaller layoff signal and thus a flatter post-layoff learning path. Note that, under additional asymmetric information assumptions that we may be concerned are at play in the background of the implications discussed in Section 1.4, this equation can be modified to include the interaction of $AFQT^*$ and tn_{it} (tenure on worker's current job) or the interaction of $AFQT^*$ and js_{it} (the length of a worker's current job spell), which addresses and expands upon the empirical estimation frameworks of Schönberg (2007a) and Pinkston (2009), respectively.

The key aspects of this regression model are in how it relates to the predictions about the combined return to ability following a layoff discussed in Section 1.4. In this regression model, β_{ax} represents the experience-ability profile for workers for whom $D_{it} = 0$ and is analogous to the β_{ax} term from Equation (1.6) for the sample of non-laid-off workers; β_a^0 represents the discrete change in the return to $AFQT^*$ for laid-off workers relative to non-laid-off workers; $\beta_{ax'}^0$ represents the return to the post-layoff experience-ability profile; β_a^x represents the return to the pre-layoff experience-ability profile; and $\beta_{ax'}^x$ represents the way in which pre- and post-layoff experience interact with each other and the ability measure. To see how these predictions relate to the overall return to a worker's residual AFQT score, observe that

$$
\frac{\partial w_{it}}{\partial AFQT_i^*} = \beta_a + \underbrace{\beta_{ax} Exp_{it}}_{\substack{\text{Normal Learning} \\ \text{Process}}} + D_{it}\Big[\beta_a^0 + \underbrace{\beta_a^x Pre_{it}}_{\substack{\text{Pre-Layoff} \\ \text{Learning}}} + \underbrace{\beta_{ax'}^0 Post_{it}}_{\substack{\text{Post-Layoff} \\ \text{Learning}}} + \underbrace{\beta_{ax'}^x Pre_{it} \times Post_{it}}_{\frac{\Delta \text{ Post-Layoff Learning}}{\Delta \text{ Pre-Layoff Learning}}}\Big].
$$

Following AP's omitted variables bias approach, it can be shown that

$$
\begin{aligned}
b_{sx} &= \gamma^* + \Phi_{qs} \\
b_{zx} &= \Phi_{zx} + \sum_{k \le x} D^k \left(\delta^k \Phi_{zx} + \Phi_{zk} + \Phi_{zx}^k + \Phi_z^k\right) \times \Pr\left(D^k = 1\right)
\end{aligned}
$$

The coefficient b_{zx} in this set up ends up being essentially analogous to a random coefficients model - see Wooldridge (2010) page 74 for discussion on random coefficients models.

If a layoff truly conveys a negative ability signal, the predictions discussed in Section 1.4 suggest that a layoff will disproportionately hurt high residual ability workers ($\beta_a^0 < 0$ - prediction (i)), while the return to the residual ability-experience profile will be greater for laid-off workers with both pre- and post-layoff experience ($\beta_{ax'}^0, \beta_a^x > \beta_{ax} > 0$ - predictions (ii) and (iii)), and the increased return to the residual ability-post-layoff experience profile will decrease in pre-layoff experience ($\beta_{ax'}^x < 0$ - prediction (iv)).

1.6 Empirical Model Estimation

Table 1.5 reports estimates based on the predictions regarding the signaling role of layoffs discussed in Section 1.4. Column 1 reports the traditional employer learning model estimates based on Equation (1.6), while Column 2 reports the same model with the inclusion of an ability-job loss variable. Comparing the estimated coefficients on AFQT$^* \times$ Total Exp tests the prediction that the AFQT-experience profile is different between the two estimation models. The difference between the two estimates indicates that when the AFQT-job loss interaction variable is added, the return to a standard deviation increase in residual AFQT score after 10 years of potential labor market experience increases from 0.003 log points to 0.004 log points, with both estimates significant at the 0.1% significance level. While this finding is encouraging, the model also included the interaction between residual AFQT score and an indicator for job loss due to plant closure, with the coefficients for both this and the layoff interaction very imprecisely estimated and indistinguishable from zero.

The rest of the columns in Table 1.5 illustrate the complex relationship between the layoff signal and pre- and post-layoff experience, culminating in Column 5 which estimates the main estimation model discussed above in Equation (2.8). Column 3 adds an interaction between residual AFQT score and post-layoff experience. When the post-experience interaction is added, the coefficient on the layoff-AFQT interaction is negative (-0.031 log points) and significant at the 5% significance level, however the interaction between AFQT score and post-layoff experience is not, though the sign is correct. The bigger issue with the estimates in Column 3 is that the interaction between closed and AFQT is negative (-0.028 log points), which could suggest that job loss in general hurts higher ability workers more than lower ability workers. This issue is no longer present when the interaction between pre-layoff experience and AFQT score is added to the regression model in Column 4, with all three main layoff-AFQT coefficients as predicted by the model, however they are imprecisely estimated.

Turning attention now to Column 5, which provides estimates for the coefficients from the full regression

Table 1.5: Employer Learning Around Layoff versus Plant Closure

Independent Variable	(1) Log Wage	(2) Log Wage	(3) Log Wage	(4) Log Wage	(5) Log Wage
Education	**0.125*****	**0.118*****	**0.117*****	**0.118*****	**0.121*****
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Educ × Total Exp	**-0.002**	**-0.001**	**-0.001**	**-0.001**	**-0.002**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
AFQT*	**-0.003**	**-0.004**	**-0.003**	**-0.000**	**-0.000**
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
AFQT* × Total Exp	**0.003*****	**0.004*****	**0.004*****	**0.004****	**0.004****
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Layoff × AFQT*		**-0.021**	**-0.031***	**-0.047***	**-0.075****
		(0.013)	(0.014)	(0.023)	(0.023)
Layoff × AFQT* × Post Exp			**0.002**	**0.007***	**0.014****
			(0.003)	(0.003)	(0.005)
Layoff × AFQT* × Pre Exp				**0.004**	**0.012****
				(0.004)	(0.004)
Layoff × AFQT* × Post × Pre Exp					**-0.003***
					(0.001)
Closed × AFQT*		**-0.000**	**-0.028**	**0.031**	**0.010**
		(0.021)	(0.024)	(0.043)	(0.045)
Closed × AFQT* × Post Exp			**0.006**	**0.009***	**0.011**
			(0.004)	(0.004)	(0.010)
Closed × AFQT* × Pre Exp				**-0.013**$^{+}$	**-0.007**
				(0.008)	(0.008)
Closed × AFQT* × Post × Pre Exp					**-0.001**
					(0.002)
R^2	0.369	0.381	0.383	0.385	0.387
Observations	139,146	139,146	139,146	139,146	139,146
Individuals	3,653	3,653	3,653	3,653	3,653
No. of Layoffs	.	711	711	711	711
(Avg. Year)	(.)	(2007.35)	(2007.35)	(2007.35)	(2007.35)
No. of Plant Closings	.	206	206	206	206
(Avg. Year)	(.)	(2007)	(2007)	(2007)	(2007)

Cluster-robust standard errors in parentheses are computed at the individual worker level
$^{+}$ $p < 0.10$, $^{++}$ $p < 0.075$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: The first year of potential experience is dropped from the analysis as this was used to create the residual ability measures. All models include a quadratic in pre- and post- job loss potential experience, indicators for each type of job loss, pre-job loss experience interacted with post-job loss experience, indicators for if the job loss took place between 2008 and 2010 or after 2010, a vector of year-quarter indicators, education interacted with a vector of year indicators, indicators for race, female, union status, part-time status, two-digit entry industry, and the log number of employees at the workers entry job, all interacted with a cubic time trend. The base year for the year indicators and time trends is 2017. Additionally, with the exception of the two-digit entry industry dummies, each individual control above is also interacted with a quadratic in pre- and post-job loss potential experience.

model in Equation (2.8). Under this specification, the interaction between layoff and AFQT is negative (-0.075 log points) and significant, while the pre-layoff experience-AFQT interaction is positive (0.012 log points) and significant at the 1% level. Similarly, the post-layoff experience-AFQT interaction is also positive (0.014 log points) and significant (at the 1% level), and it is estimated precisely enough to be statistically different from the estimated AFQT-exp interaction for the non-job loser sample. Additionally, the interaction between AFQT score and pre- and post-layoff experience is negative (-0.003 log points) and significant (at the 5% level). All of these coefficients match the predictions of the learning model developed in this paper. A similar result is not found for the sample of worker who lost their jobs due to plant closure, which suggests that the estimated coefficients from this regression provide strong evidence in support of the signaling role of layoffs.

In literal terms, the estimates in Column 5 in Table 1.5 indicate that among the sample of laid-off workers, a standard deviation increase in residual AFQT score decreases wages following a layoff by roughly seven percent, with this effect decreasing by around one percentage point per year of pre-layoff experience (or 12 percentage points over 10 years). Following a layoff, a standard deviation increase in residual AFQT score is associated with a nearly 14 percent increase in wages 10 years after the event, with this effect decreasing by around 3 percentage points per year of pre-layoff experience. These estimates generally back up the predictions of the conceptual framework discussed previously, and lend support to the idea that signaling is playing a role in the overall costs of job loss for young laid-off workers, especially during the first five years or so of labor market experience.

1.6.1 Results For Alternate Sample

Table 1.6 provides estimation results based on Equation (2.8) for different samples and control specifications. Column 1 shows the results from the main specification discussed above. As discussed previously, changing education level may bias the estimated ability-experience profile for the total sample, and this same logic holds for the estimated effects of the layoff-ability interactions. Thus, Column 2 repeats the main specification above while excluding observations for workers who change education levels beginning two years prior to the reported change to account for decreased labor market participation due to re-enrollment during the period. While less precise and slightly smaller, the results in Column 2 on quantitatively similar to those found in the main specification. As there may be concerns that these results are being driven by high-ability workers losing their jobs during the Great Recession, Columns 3 and 4 of Table 1.6 repeat the specifications of Columns 1 and 2 respectively, and again report quantitatively similar results.[29] Finally, as the main specification allowed for observable characteristics to interact with the job loss experience profile, Columns 5-9 repeat the estimation of the models in Columns 1-4 without including the interaction controls. Again, these results are quantitatively similar to the preferred specification in Column 1 and generally significant. Taken together, the general consistency of the results across model specifications and samples reinforces the notion that the layoff-ability effects are driven by the signaling nature of layoffs.

1.6.2 Differences By Race and Gender

The empirical evidence based on the approach developed in Section 1.5 strongly suggests that layoff signaling is occurring in the labor market for young workers at the start of their careers. While this is an important finding for understanding the consequences of job loss for young workers as whole, the evidence is based on the fairly strong assumption that employers learn about all types of workers in the same way and at the same speed. That is, like the FG and AP models, the approach developed in the previous section assumes that the way that employers learn about a worker's ability is independent of observable characteristics, such as race, gender, or education. While this assumption simplifies the analysis, a number of studies have documented differences in employer learning across groups defined along various dimensions, thus violating the assumption.[30] In this subsection, I modify the employer learning estimation strategies described by Equations (1.6) and (2.8) to allow the return to a worker's AFQT score with experience to vary based on group membership, defined by race, gender, or education, in order to investigate whether differences in learning across groups yield heterogeneous layoff signaling effects.

Table 1.7 provides estimated effects when allowing learning effects to vary by gender. Column 1 shows shows the results based on the inclusion of the interaction of an indicator for female and residual AFQT score into the regression model described by Equation (1.6). Based on this baseline model, female workers have a higher return to ability with experience than male workers, totaling around three percent over 10 years for a standard deviation increase in residual AFQT. Columns 2 and 3 show estimated effects from

[29] See H. Farber (2017) for more information on the differences in the impact of involuntary job loss prior to and after the Great Recession.

[30] For instance Arcidiacono et al. (2010) and Light and McGee (2015a) document differences in employer learning across different education groups; Mansour (2012) documents differences in learning across initial occupations; Pinkston (2003) and Castex and Dechter (2014) document differences across gender; and Pinkston (2006) and Arcidiacono et al. (2010) document differences by race.

Table 1.6: Employer Learning Around Layoff versus Plant Closure — Robustness Checks

Independent Variable	Log Wage	Log Wage	Log Wage	Log Wage	Log Wage	Log Wage	Log Wage	Log Wage
Education	**0.121*****	**0.130*****	**0.118*****	**0.127*****	**0.127*****	**0.136*****	**0.127*****	**0.144*****
	(0.022)	(0.024)	(0.023)	(0.026)	(0.022)	(0.025)	(0.022)	(0.025)
Educ × Total Exp	**-0.002**	**-0.003**	**-0.002**	**-0.003**	**-0.002**	**-0.003**$^{++}$	**-0.002**	**-0.004***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
AFQT*	**-0.000**	**-0.000**	**-0.000**	**0.000**	**-0.001**	**-0.000**	**-0.000**	**0.000**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
AFQT* × Total Exp	**0.004****	**0.004****	**0.003****	**0.004****	**0.004*****	**0.004****	**0.004****	**0.004****
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Layoff × AFQT*	**-0.075****	**-0.076****	**-0.085***	**-0.078***	**-0.068****	**-0.067****	**-0.072**$^{++}$	**-0.068**$^{+}$
	(0.023)	(0.024)	(0.038)	(0.039)	(0.023)	(0.024)	(0.037)	(0.041)
Layoff × AFQT* × Post Exp	**0.014****	**0.015****	**0.018****	**0.019****	**0.014****	**0.014****	**0.016****	**0.017***
	(0.005)	(0.005)	(0.006)	(0.007)	(0.004)	(0.005)	(0.006)	(0.007)
Layoff × AFQT* × Pre Exp	**0.012****	**0.012***	**0.014**	**0.011**	**0.009***	**0.009**$^{++}$	**0.008**	**0.006**
	(0.004)	(0.005)	(0.015)	(0.016)	(0.004)	(0.005)	(0.014)	(0.016)
Layoff × AFQT* × Post × Pre Exp	**-0.003***	**-0.003***	**-0.005**$^{+}$	**-0.005**	**-0.002**$^{+}$	**-0.003**$^{++}$	**-0.003**	**-0.004**
	(0.001)	(0.002)	(0.003)	(0.003)	(0.001)	(0.002)	(0.003)	(0.003)
Closed × AFQT*	**0.010**	**-0.012**	**0.011**	**-0.021**	**-0.003**	**-0.011**	**0.015**	**0.037**
	(0.045)	(0.051)	(0.068)	(0.070)	(0.050)	(0.056)	(0.061)	(0.068)
Closed × AFQT* × Post Exp	**0.011**	**0.010**	**0.010**	**0.018**	**0.018**	**0.013**	**0.014**	**0.016**
	(0.010)	(0.017)	(0.013)	(0.023)	(0.011)	(0.017)	(0.013)	(0.023)
Closed × AFQT* × Pre Exp	**-0.007**	**-0.002**	**-0.005**	**0.009**	**-0.002**	**0.001**	**-0.001**	**-0.007**
	(0.008)	(0.009)	(0.018)	(0.020)	(0.009)	(0.010)	(0.015)	(0.017)
Closed × AFQT* × Post × Pre Exp	**-0.001**	**0.001**	**-0.001**	**-0.003**	**-0.003**	**-0.001**	**-0.002**	**-0.002**
	(0.002)	(0.004)	(0.003)	(0.006)	(0.003)	(0.004)	(0.004)	(0.006)
Fired × AFQT*	**-0.052**$^{+}$	**-0.040**	**0.014**	**0.021**	**-0.045**	**-0.039**	**0.010**	**0.006**
	(0.030)	(0.033)	(0.039)	(0.042)	(0.028)	(0.033)	(0.037)	(0.043)
Fired × AFQT* × Post Exp	**0.003**	**0.011**$^{++}$	**0.002**	**0.010**	**0.002**	**0.008**	**-0.001**	**0.006**
	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)
Fired × AFQT* × Pre Exp	**0.003**	**0.001**	**-0.015**	**-0.019**	**0.002**	**0.001**	**-0.014**	**-0.015**
	(0.006)	(0.007)	(0.012)	(0.014)	(0.005)	(0.007)	(0.012)	(0.014)
Fired × AFQT* × Post × Pre Exp	**0.001**	**-0.002**	**0.000**	**-0.002**	**0.001**	**-0.002**	**0.001**	**-0.001**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Controls for Interactions	Yes	Yes	Yes	Yes	No	No	No	No
Drops if Education Changes	No	Yes	No	Yes	No	Yes	No	Yes
Removes Job Losses After 2007	No	No	Yes	Yes	No	No	Yes	Yes
R^2	0.387	0.400	0.394	0.407	0.378	0.389	0.387	0.396
Observations	139,146	103,893	127,185	95,438	139,146	103,893	127,185	95,438
Individuals	3,653	3,396	3,650	3,393	3,653	3,396	3,650	3,393
No. of Layoffs	711	579	326	284	711	579	326	284
(Avg. Year)	(2007.35)	(2007.11)	(2004.09)	(2004.12)	(2007.35)	(2007.11)	(2004.09)	(2004.12)
No. of Plant Closings	206	163	111	94	206	163	111	94
(Avg. Year)	(2007.09)	(2006.88)	(2004.35)	(2004.41)	(2007.09)	(2006.88)	(2004.35)	(2004.41)
No. of Firings	495	398	280	239	495	398	280	239
(Avg. Year)	(2006.64)	(2006.34)	(2004.19)	(2004.12)	(2006.64)	(2006.34)	(2004.19)	(2004.12)

Cluster-robust standard errors in parentheses
$^{+}$ $p < 0.10$, $^{++}$ $p < 0.075$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors are clustered at the individual worker level. The first year of potential experience is dropped from the analysis as this was used to create the residual ability measures. Columns 1-4 report results for estimation that include each individual control from Table 1.5 interacted with a quadratic in pre- and post-job loss potential experience, while columns 5-8 do not include these controls. Columns 2, 4, 6, and 8 drop respondents two years prior to any education level changes. entering the labor market. Columns 3, 4, 7, and 8 drop respondents who lose their job for the first time in 2008 or later.

an extension of Equation (2.8) that includes the layoff ability variables interacted with a female indicator. Coefficients suggest that female workers experience smaller layoff signaling effects than male workers, as the estimates on the interacted variables suggest an overall lower affect when compared with the coefficients on the reference variables.

Table 1.7: Log Wage Regressions Using Potential Experience — By Gender

	(I)	(II)	(III)
AFQT*	**-0.000**	**0.001**	**-0.007**
	(0.007)	(0.007)	(0.009)
AFQT* × Exp	**0.002**$^{+}$	**0.002**	**0.003***
	(0.001)	(0.001)	(0.002)
Layoff × AFQT*	**-0.068****	**-0.071****	**-0.064***
	(0.023)	(0.027)	(0.028)
Layoff × AFQT* × Post Exp	**0.014****	**0.015****	**0.015****
	(0.004)	(0.005)	(0.005)
Layoff × AFQT* × Pre Exp	**0.012****	**0.012***	**0.012***
	(0.004)	(0.005)	(0.005)
Layoff × AFQT* × Post × Pre Exp	**-0.002**$^{+}$	**-0.003***	**-0.003***
	(0.001)	(0.001)	(0.001)
Female × AFQT*			**0.017**
			(0.014)
Female × AFQT* × Exp	**0.003**$^{++}$	**0.003**$^{++}$	**0.001**
	(0.002)	(0.002)	(0.002)
Female × Layoff × AFQT*		**0.035**	**0.018**
		(0.049)	(0.050)
Female × Layoff × AFQT* × Post Exp		**-0.005**	**-0.005**
		(0.010)	(0.010)
Female × Layoff × AFQT* × Pre Exp		**-0.002**	**-0.002**
		(0.009)	(0.009)
Female × Layoff × AFQT* × Post × Pre Exp		**0.004**	**0.004**
		(0.002)	(0.002)
Observations	139,146	139,146	139,146
Individuals	3,653	3,653	3,653
Unique Males	.	1,830	1,830
No. of Male Layoffs	.	458	458
Unique Females	.	1,823	1,823
No. of Female Layoffs	.	253	253

Cluster-robust standard errors in parentheses are computed at the individual worker level

$^{+}$ $p < 0.10$, $^{++}$ $p < 0.075$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.8 provides estimated effects when allowing learning effects to vary by race. Column 1 shows shows the results based on the inclusion of the interaction of indicators for Black and Hispanic and residual AFQT score into the regression model described by Equation (1.6). Based on this baseline model, Black workers have a higher return to ability with experience than white or Hispanic workers, totaling around three percent over 10 years for a standard deviation increase in residual AFQT. Columns 2 and 3 show estimated effects from an extension of Equation (2.8) that includes the layoff ability variables interacted with indicators for Black and Hispanic. Coefficients suggest that, unlike white and Hispanic workers, Black workers experience no noticeable layoff signal effect, as the estimates on the interacted variables almost completely offset the estimates from the reference variables.

Table 1.8: Log Wage Regressions Using Potential Experience — By Race

	(I)	(II)	(III)
AFQT*	**-0.001**	**-0.001**	**-0.010**
	(0.007)	(0.007)	(0.010)
AFQT* × Exp	**0.003***	**0.003***	**0.004****
	(0.001)	(0.001)	(0.002)
Layoff × AFQT*	**-0.068****	**-0.115*****	**-0.107*****
	(0.023)	(0.031)	(0.032)
Layoff × AFQT* × Post Exp	**0.014****	**0.017****	**0.017****
	(0.004)	(0.006)	(0.006)
Layoff × AFQT* × Pre Exp	**0.012****	**0.017****	**0.017****
	(0.004)	(0.006)	(0.006)
Layoff × AFQT* × Post × Pre Exp	**-0.002**$^{+}$	**-0.003**$^{+}$	**-0.003**$^{+}$
	(0.001)	(0.002)	(0.002)
Black × AFQT*			**0.016**
			(0.017)
Black × AFQT* × Exp	**0.004***	**0.004***	**0.002**
	(0.002)	(0.002)	(0.003)
Black × Layoff × AFQT*		**0.127***	**0.111***
		(0.050)	(0.053)
Black × Layoff × AFQT* × Post Exp		**-0.005**	**-0.005**
		(0.011)	(0.011)
Black × Layoff × AFQT* × Pre Exp		**-0.011**	**-0.011**
		(0.010)	(0.010)
Black × Layoff × AFQT* × Post × Pre Exp		**0.002**	**0.002**
		(0.003)	(0.003)
Hispanic × AFQT*			**0.021**
			(0.017)
Hispanic × AFQT* × Exp	**-0.003**	**-0.003**	**-0.005***
	(0.002)	(0.002)	(0.003)
Hispanic × Layoff × AFQT*		**0.063**	**0.042**
		(0.050)	(0.053)
Hispanic × Layoff × AFQT* × Post Exp		**-0.006**	**-0.006**
		(0.010)	(0.010)
Hispanic × Layoff × AFQT* × Pre Exp		**-0.007**	**-0.007**
		(0.009)	(0.009)
Hispanic × Layoff × AFQT* × Post × Pre Exp		**0.001**	**0.001**
		(0.003)	(0.003)
Observations	139,146	139,146	139,146
Individuals	3,653	3,653	3,653
No. of White Layoffs	.	327	327
No. of Black Layoffs	.	228	228
No. of Hispanic Layoffs	.	156	156

Cluster-robust standard errors in parentheses are computed at the individual worker level
$^{+}$ $p < 0.10$, $^{++}$ $p < 0.075$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

1.7 Conclusion

In this paper, I examine the extent to which ability signaling explains the long-term wage losses suffered by young workers who experience layoffs. Consistent with the notion that employers use layoffs as signals of productive ability (Gibbons and Katz 1991; Michaud 2018), I find fairly strong evidence in support of ability signaling following layoffs but not plant closures. As predicted by the model, ability signals play a relatively large role in determining wages early in a worker's labor market career when information asymmetry is greatest. Results suggest that layoff signals allow employers to adjust wages of low relative ability workers downward more quickly, driving persistent earning losses. High relative ability workers' wages initially decrease following job loss but recover as post-layoff experience increases.

As in previous work on employer learning (*e.g.* Pinkston 2003; Arcidiacono et al. 2010), I find heterogeneous effects of layoff signaling by race and gender. Future work may further examine drivers of these differences and their role in explaining racial and gender wage gaps. Additionally, researchers may study the extent to which effects of layoff signaling vary by workers' levels of specific skills, such as math, verbal, and social skills.

Combined with evidence from previous literature (Lachowska et al. 2020), the results that I find highlight differences in the sources of earnings losses across displaced workers with different levels of experience. Whereas researchers find that loss of firm-specific human capital and high-quality job matches drive earnings losses following job loss among older workers (cite), the signaling nature of layoffs is unique to young workers and warrants a separate set of policy solutions. For example, policies that encourage positive ability signaling, such as returning to school to obtain a GED or master's degree, likely will be more effective when targeted at younger workers. Nonetheless, similarly to older workers, younger workers of all ability levels tend to experience substantial decreases in wages following job loss. Thus, workers of all ages may benefit most from programs and services that prevent workers from experiencing layoffs in the first place

Finally, while the primary focus of the paper is on understanding the sources of earnings losses associated with early career job loss, the results are in line with previous literature that empirically identifies asymmetric employer learning more generally (Pinkston 2009; L. B. Kahn 2013). The methodology I use in this study, as well as the focus on how the information on worker ability evolves with experience is fairly novel in the asymmetric employer learning literature. This approach, coupled with the formal extension of AP's employer learning model pursued in VanderBerg (2021a), should provide additional avenues for future research on the effects of asymmetric employer learning.

CHAPTER 2

EMPLOYER LEARNING AND STATISTICAL DISCRIMINATION WITH UNEXPECTED INFORMATION

2.1 Introduction

When a worker first enters the labor market, employers lack full information about their true productive ability. As that worker gains labor market experience, however, employers gradually learn about their ability and set wages accordingly, yielding higher returns to ability among more experienced workers (H. Farber and Gibbons 1996; Altonji and Pierret 2001). In existing work, H. Farber and Gibbons (1996) and Altonji and Pierret (2001)—hereafter, FG and AP, respectively—model the revelation of workers' ability to employers. FG and AP's models operate under the assumptions that employer learning is a smooth, continuous process and workers' observable characteristics are fixed at labor market entry. In practice, though, many workers' observable characteristics change discretely over the course of their labor market careers. For example, about one-third of men who first left school between 1978 and 1990 had returned by 1991 (Light 1995). If increases in educational attainment or other mid-career changes in workers' observable characteristics signal ability to prospective employers (Spence 1973), then they could generate discontinuities in the employer learning process.[1]

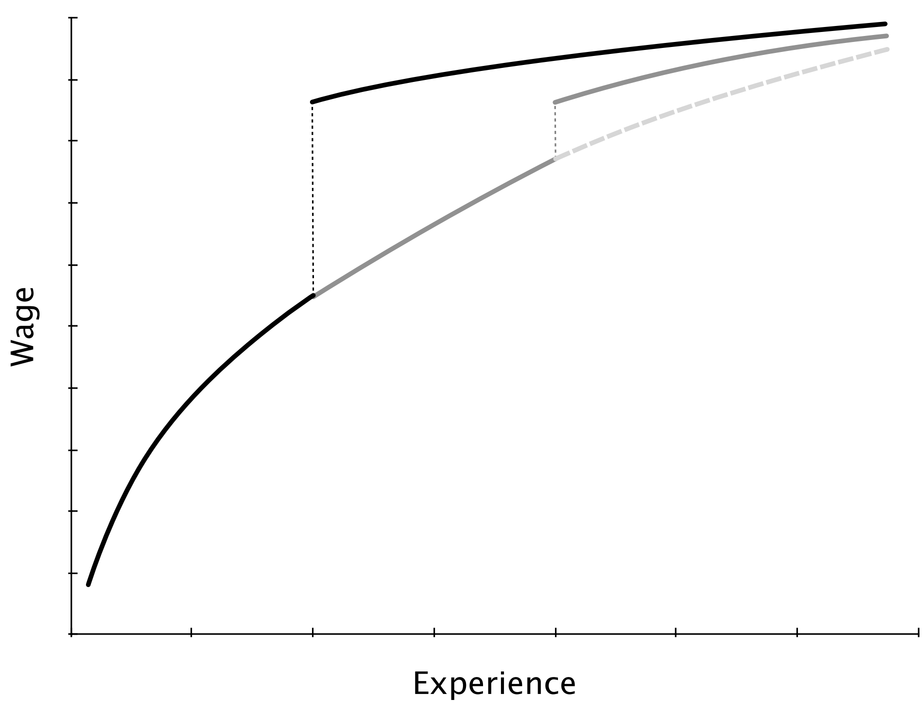
In this paper, I model employer learning in the presence of unexpected changes in observable characteristics, which I refer to as "events," and allow for them to interrupt the smooth, continuous employer learning processes modeled by FG and AP. Specifically, I generalize their models to allow employers to update prior beliefs about workers and to set wages based on post-labor market entry events, such as returning to school or experiencing a layoff. My model allows the extent to which wages change following an event to depend on the amount of employer learning that occurred before the event time. In other words, because employers learn about workers' abilities over time, there is more capacity for a signaling role of, say, education among workers with two years of labor market experience than among workers with twenty years of labor market experience. Consistent with this notion, I show that the more labor market experience a worker has prior to

[1]This is further discussed by Light and McGee (2015a) who replicate Arcidiacono et al. (2010)'s work on the differences in employer learning by educational attainment to highlight the sizable difference in the estimated effect of the ability-experience profile when education is allowed to increase after entry versus when workers are dropped from the sample when their education levels change.

an event, the less employers infer about a worker's ability as a result of the event.

The intuition behind the wage growth predicted by this generalized employer learning model in the context of returning to school is shown in Figure 2.1, for a worker with higher than expected ability at labor market entry. If the worker returns to school early in their labor market career (solid black line), employers rely heavily on that observation to infer the worker's ability, and thus there is a large, discrete jump in the worker's wage. If the worker returns to school later in their labor market career (solid gray line), however, the wage change is smaller as employers have already learned something about the worker's ability and factored it into the worker's wage. In other words, the wage return to the information gained by employers due to the worker returning to school diminishes as the worker's labor market experience increases. Figure 2.1 also highlights that wages grow slower with experience after returning to school than they would have grown had the worker never returned to school (dashed gray line). Essentially, the additional information associated with the worker returning to school causes wage gains due to employer learning to be front-loaded relative to the smooth, gradual wage gains associated with employer learning in the absence of returning to school.

Figure 2.1: Wage Growth Under Continuous Employer Learning and After Returning to School



Notes: Example wage growth paths due to employer learning for a worker with higher than expected ability at labor market entry. The continuous line (solid black to gray to dated light gray) represents the worker's wage growth path under the continuous employer learning assumption of the EL-SD model. The discontinuous solid black and light gray lines show how wages grow due to discrete jumps in employer learning if the worker returns to school early in their labor market career.

I use data from the National Longitudinal Survey of Youth 1997 (NLSY97) to formally test my model in the context of the signaling role of returning to school.[2] While there exists a vast literature on the signaling role of educational attainment before labor market entry, I am one of the first to examine the signaling role of schooling among those who have participated in the labor market for a period of time.[3] I find suggestive evidence that returning to school to receive a GED or graduate degree sends a positive ability signal to the labor market.

The rest of this paper is organized as follows. In Section 2.2 I discuss background information and review related literature. In Section 2.3 I build my employer learning model and compare the predictions to that of the traditional employer learning model. In Section 2.4 I discuss the data used in my primary analysis and provide sample statistics of key variables. In Section 2.5 I discuss and empirically test the predictions of my model related to the signaling nature of returning to school. Section 2.6 concludes.

2.2 Background Information and Related Literature

This paper relates to the large body of literature on empirically identifying employer learning in general, while also contributing to the education signaling literature. In what follows, I first provide a general review of the existing literature on employer learning under symmetric and asymmetric information before briefly reviewing the literature on education signaling, with a specific focus on the signaling nature of returning to school.

2.2.1 Empirical Employer-Learning

The traditional empirical employer learning models developed by FG and AP, which are built off of the symmetric information model of Harris and Holmstrom (1982), assume that employers base their initial beliefs about a worker's true ability around a set of easy to observe characteristics that are time-invariant (e.g. education), while updating their beliefs relative to noisy output signals as the worker's labor market experience increases. Specifically, they show that ability correlates that are available to the researcher but

[2]Like FG and AP, I rely on AFQT scores–a general aptitude measure unique to NLS data–to serve as a measure of ability that is unobserved by employers. Other work that uses these scores as ability measures to study employer learning include Schönberg (2007a), Lange (2007), Pinkston (2009), Arcidiacono et al. (2010), Mansour (2012), and Light and McGee (2015a), among others.

[3]Examples of papers on the signaling role of education at labor market entry include Altonji and Pierret (1998), Arcidiacono et al. (2010), and Light and McGee (2015a), among others. To my knowledge, there is limited research that specifically focuses on the signaling nature of returning to school. Two notable exceptions to this are (Tyler, Murnane, and Willett 2000), who find evidence of the GED being used as a signal of ability, and Hussey (2012), who finds that a sizable portion of the return to an MBA degree is the result of signaling/screening.

not prospective employers, should not play a major factor in determining a worker's early career wages. Then, as a worker's labor market experience increases, productivity signals allow employers to learn about the worker's true ability, which increases the role of these ability correlates in determining a worker's wage; and, in the case of AP's EL-SD model, decreases the role that easy to observed measures play in determining a worker's wage.

Both FG and AP test their models using data from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79), specifically each worker's Armed Forces Qualifying Test (AFQT) score, which is a measure of an individual's general aptitude and is reasonably assumed to be unobserved by a worker's employer.[4,5] They show that the coefficient on the interaction between a worker's AFQT score and potential experience is positive and significant, illustrating that higher ability, as proxied by AFQT score, is rewarded more as experience grows and employer's learn about the worker; conversely, the coefficient on the AFQT score variable alone is close to zero, illustrating that high unobserved ability is not initially rewarded by employers. AP also show that when the interaction between AFQT score and experience is included in a log wage regression along with the interaction between years of schooling and experience, the coefficient on the latter is negative, indicating that employers are shifting weight off of the easily observed education level and onto the hard to observe ability measure as they learn more about the worker.

Other authors have expanded the general employer learning model in a number of ways. Using the same sample as AP, Lange (2007) modifies AP's EL-SD model to allow for the speed of employer learning to be structurally identified and finds that employers reduce their initial expectations errors by 50% by the time workers reach three years of potential experience. Mansour (2012) expands on the AP model to test for differences in employer learning across initial occupation and finds substantial heterogeneity in learning across initial occupation independent of a worker's education. Similarly, Arcidiacono et al. (2010) break AP's NLSY79 sample into two separate samples based on highest education level attained (high school and college), and find that employer learning is only present for non-college graduates. Light and McGee (2015a), however, show that Arcidiacono et al. (2010)'s results are quite sensitive to how labor market entry is defined, and they find evidence that employer learning does occur for college graduates. Light and McGee (2015b) and Petre (2018) break from using AFQT scores alone as their measure of hard to observe

[4] AFQT scores are derived from components of the Armed Services Vocational Aptitude Battery (ASVAB), a ten-part general skills and aptitude test administered by the United States military to gauge an individual's proficiency along a number of dimensions (mathematical, verbal reasoning, etc.).

[5] See Lange (2007) for a discussion of the assumptions surrounding the use of AFQT scores as an unobserved measure of a worker's ability.

ability and use FG and AP's models, respectively, to investigate the importance of different skill dimensions (ASVAB component test scores) and ability types (cognitive versus non-cognitive) on employer learning and find evidence of differential learning across different skills and abilities.[6]

The common driving force behind these applications of FG and AP's employer learning models is the rather strong assumption that information on workers' ability is learned symmetrically across all employers. Relaxing this assumption leads to the notion that employer learning is asymmetric, an idea that has a strong theoretical foundation based on the work of Waldman (1984) and Greenwald (1986) among others.[7] Recent work has attempted to empirically identify asymmetric employer learning in the context of the FG and AP models, but the results have been mixed. Schönberg (2007a) develops a two-period theoretical model based on Gibbons and Katz (1991) and derives predictions that she takes to an employer learning model based on AP. One of the key predictions of her model is that asymmetric employer learning can be identified by examining the difference between the coefficients on the interactions of AFQT score and experience and AFQT score and tenure at the current job. Based on her predictions, she finds little evidence of asymmetric employer learning, except possibly for college graduates.[8] While her model provides a convenient test for asymmetric employer learning, the reliance on job tenure in an AP style framework is prohibitive, especially if information is transmitted differently across employers when a worker switches jobs due to the nature of the job change, an issue that the model in this study avoids by abstracting away from the effects of learning based on tenure.

On the other hand, Pinkston (2009), in a paper closely related to this study, develops an empirical asymmetric employer learning model in which asymmetric information is passed between employers in a worker's current employment spell, as opposed to being specific to a unique employer. By looking at the difference in the coefficients on the interactions between AFQT and experience and AFQT and current employment spell, he finds that asymmetric learning has at least as large of an effect on wages as public learning during an employment spell.[9] While Pinkston (2009)'s model allows for the identification of asymmetric employer

[6]See Speer (2017a) for a discussion regarding the ASVAB component tests and other types of cognitive and non-cognitive ability, and Altonji et al. (2012a) for information on the construction of AFQT scores using the ASVAB component test data in the NLSYs.

[7]See Waldman (2012) for a review of the theoretical and empirical literature that evolved out of these early models of asymmetric information.

[8]Zhang (2007) extends Schönberg (2007a)'s theoretical model to three periods and finds evidence in support of asymmetric employer learning based on the predictions of his model.

[9]Additional empirical tests for asymmetric employer learning include Devaro and Waldman (2012), L. B. Kahn (2013), and Michaud (2018), all of which find evidence in support of asymmetric employer learning, but do so outside of the FG and AP framework. Additionally, while not specifically focused on asymmetric learning, the results of Mansour (2012) suggest that asymmetric employer learning exists at least between employers across occupations.

learning in general, it does not specifically provide a means of addressing certain predictions from the literature regarding signals that are predicted to arise under asymmetric information, such as the signaling role of layoffs or attaining additional education, an issue that is addressed by the model developed in this study,

2.2.2 Returning to School to Signal Ability

One particular signaling-type event that could affect employer learning and be observable given the model developed in this paper regards workers who return to school in order to increase their education level. While not often addressed in detail, returning to school after a period in the labor market is a fairly common phenomenon (Light and McGee 2015a). In the data used in the empirical section of the paper, for instance, a little over one quarter of the sample increase their education level more than two years after they first enter the labor market. Not only does increasing education violate the traditional employer learning assumption that the easily observable characteristic vector is time-invariant, it is also plausible that workers use increased education to signal prospective employers that their ability is higher than the employer believes it is based solely on output signals. Thus, if certain types of workers are more likely than others to return to school to seek additional education, prospective employers may be able to use that as a signal of a worker's true ability that was not already revealed to them through output signals prior to returning to school.

Given the costs of returning to school (both in forgone wages and/or leisure), it is plausible to assume that workers who return to school may be doing so as a way of signaling the market that they are of higher unobserved ability than their entry education level peers, in a manner similar to Spence (1973).[10] That is, relative to their peers of the same labor market entry education level, it is plausible that the workers most likely to seek additional education will be those who are of higher residual ability.[11] Indeed summary statistics related to this group of workers identified in the data used in this survey indicate an average residual AFQT score that is around one fifth of a standard deviation higher than the mean. This raises an important issue regarding the use of education in employer learning models: if returning to education can be used a

[10]See Riley (2001) and Spence (2002) for reviews of the early education signaling literature that evolved in out of Spence's original model.

[11]This argument does raise the question of why these workers left school earlier than may have been ideal given their ability. The literature on college dropout decisions may provide some answers to this question as R. Stinebrickner and T. Stinebrickner (2012) find that bad experiences during the first couple years of college significantly increases the likelihood that a student drops out of college. If some of these bad experiences are not the result of the student learning that they are of lower ability, it seems plausible that some portion of these students will return to school at a later point and complete education more in line with their true ability. Other possible reasons for a worker to leave school early relate to inadequate incentives for high ability types to use school as a signal of ability (as in Frazis (2002), or institutional factors, such as lack of access to college (as in Bedard (2001)).

signal of a worker's ability, then the education variable used in empirical models of employer learning must be based on a worker's education level when they first entered the labor market, and cannot change over the course of a labor market career.

In addition to requiring the use of entry level education in place of current education level, different levels of additional schooling may send different ability signals to prospective employers if employers believe that different types of degrees convey different signals of ability, thus necessitating the need to specifically identify the different possible levels of education worker's can achieve by returning to school. For instance, a worker with a high school education level at labor market entry who then returns to school and gains a 2 year (or 4 year) college degree, likely is sending a different signal than a worker who entered the market as a four year college graduate who then returns to school to obtain a masters/PhD degree. Additionally, it may well be the case that different levels of post-labor market entry may be more likely to affect wages through human capital accumulation (as in Arcidiacono et al. (2016)), as opposed to ability signaling. While human capital accumulation is likely to be present for all of these workers, as long as the role that this additional human capital plays in determining worker productivity is uncorrelated with residual ability, my model will still be able to identify any signaling that may be present in the effect of returning to school on wages.

2.3 An Employer Learning Model with Unexpected Information

This section outlines the key assumptions, intuition, and empirical predictions of an augmented employer learning model that allows for unexpected post-labor market entry "events" (*e.g.* increased education) to act as signals of a worker's ability under asymmetric information. As with much of the early literature on employer learning, this model builds off of the seminal models developed by FG and AP, relying on much of the same general intuition in the setup of the model, albeit in a dynamic setting. Specifically, by incorporating assumptions from FG, the model can be seen as an augmented version of the AP model in which unexpected event signals give rise to a form of statistical discrimination throughout a worker's labor market career. Given the similarity of the initial setup of the model to the general FG and AP framework, the primary focus of this section is to present a general overview of the model setup, derivations, intuition, and results in a succinct and tractable manner. As such, I relegate the formal presentation of the FG/AP model to Appendix B.1, and the formal presentation of the benchmark model discussed below to Appendix B.2, where I formally derive the predictions discussed here.

2.3.1 Basic Setup

Following the AP employer learning structure, assume that each worker in the labor market is endowed with some level of time-invariant productive ability that is not directly observed by employers in the labor market or by researchers. Suppose, however, that researchers have access to an ability correlate for each worker (say a test score) that serves as an imperfect proxy for the worker's total ability, and is not directly observed by employers. This allows a worker's total productive ability to be expressed as $\beta z + \eta$, which is the sum of a function of the observed ability correlate (z) and the remaining portion of the worker's ability (η) that is orthogonal to the observed ability correlate.[12] Each worker also enters the labor market with time-invariant characteristics that directly affect productivity and are easily observed by all employers, but are only partially observed by researchers who have access to certain characteristics (s), such as years of general schooling, but not others (q), such as specific educational skill programs. Additionally, unlike in the AP model, after a worker enters the labor market, it is possible for them to experience some sort of event (d), which may have a direct effect on the worker's productivity. It is assumed that this event is observed by all employers in the labor market and by researchers.

In the absence of complete information regarding a worker's true productive ability when the worker first enters the labor market, employers must initially rely solely on characteristics that are easily observed (s and q) in order to form an expectation regarding the worker's true ability.[13] The resulting difference, or expectation error, between employers' initial beliefs about the worker's ability and a worker's true ability, expressed as $\beta\tilde{z} + \tilde{\eta}$,[14] leads to a discrepancy between the true value of a worker's log-productivity y and the value that employers place on the worker's expected log-productivity, $E(y|s,q)$. This discrepancy, which is a form of statistical discrimination, leads to different workers being over/under valued in entry labor markets based on their observable characteristics, rather than their true productivity. Consequently a worker's true

[12]More precisely, η represents the difference between a worker's true ability (say η^*) and the expectation of that worker's ability, conditional on z, which, for simplicity, is assumed to be linear in z. Thus, conditional on observing z, the remaining portion of a worker's unobserved ability, $\eta = \eta^* - \beta z$, is best thought of as an expectation error.

[13]Note that while labor market events (d) are easily observed, they can only occur *after* a worker has entered the labor market, and thus cannot be factored in employers' expectations over the worker's ability at the point of labor market entry.

[14]This expression comes from the relationship between a worker's true ability parameters (z and η) and employers' conditional expectation functions of these parameters, respectively defined by $z = E(z|s,q) + \tilde{z}$ and $\eta = E(\eta|s,q) + \tilde{\eta}$, where $\tilde{z}$ and $\tilde{\eta}$ are the remaining expectation errors in employers' conditional expectation functions. See AP for further discussion on these expectation functions.

productivity at any experience level $t > 0$ can be expressed as,

$$y_t = \underbrace{\overbrace{E(y_0|s,q)}^{\text{Inital expected log-productivity}} + \overbrace{(\beta\tilde{z} + \tilde{\eta})}^{\text{Employers' initial expectation error}}}_{\text{Intial log-productivity}} + \overbrace{\delta_d d_{t_0} + \tilde{H}(t)}^{\text{Experience related characteristics}} \tag{2.1}$$

which is a function of a worker's initial log-productivity, an experience profile of productivity, $\tilde{H}(t)$, and the realization of a labor market event at experience level t_0.

In addition to the easily observed characteristics s and q, employers also rely on observing a noisy measure of a worker's output at each level of experience to learn more about a worker's true ability. While this measure does not perfectly reveal y, it can be used as a noisy signal of the worker's true ability denoted, I_t, conditional on the observable characteristics s and q.[15] Then, at the beginning of the worker's next experience level, employers use these signals to update their beliefs about a worker's ability by incorporating the entire history of output signals into their new expectation. Importantly, in this basic setup, these signals are assumed to be seen by all employers in the labor market. From this, it is possible to rewrite a worker's log-productivity function to account for updated beliefs about a worker's ability based solely on output signals $D_t = (I_0, \ldots, I_{t-1})$ as,

$$y_t = \overbrace{E(y_0|s,q)}^{\text{Initial expected log-productivity}} + \overbrace{E(\beta\tilde{z} + \tilde{\eta}|D_t)}^{\text{Expected ability given } D_t} + \overbrace{(\beta\tilde{z}_t + \tilde{\eta}_t)}^{\text{Remaining error given } D_t} + \overbrace{\delta_d d_{t_0} + \tilde{H}(t)}^{\text{Experience related characteristics}} \tag{2.2}$$

which, ignoring the effect of d, leads to the log-wage function described in AP, where wage changes occur due to actual changes in productivity related to $\tilde{H}(t)$ and as a result of employers' correctly adjusting wages to match the worker's expected ability

Finally, unlike in previous employer learning models, in addition to the output signals observed at each experience level, assume that employers also share the common belief that observing that a worker has experienced a labor market event (d) reveals information about the worker's true ability. In the next subsection, I will show that when signaling events are the result of asymmetric information between workers and employers, the relationship between an event signal and experience essentially results in a dynamic version of AP's statistical discrimination model.

[15]Following AP, this signal can more formally be expressed as $I_t = \beta\tilde{z} + \tilde{\eta} + \nu_t$, where ν_t represents transitory variation in a worker's output and the firm's production characteristics that are not easily accounted for when evaluating a worker.

2.3.2 Information Structure

While recent evidence suggests that employer learning is asymmetric (Pinkston 2009; L. B. Kahn 2013), information in this model is assumed to be symmetric across employers. This assumption greatly simplifies the derivations of the model and provides a convenient benchmark for analyzing the signaling nature of certain labor market events, even those associated with asymmetric information.[16] Further, the symmetric learning assumption yields a model that is directly comparable to the early employer learning models developed out of FG and AP. The information structure based on the setup above under symmetric information across employers is summarized as follows:[17]

1. The i^{th} worker's total productive ability $\beta z_i + \eta_i$, is given by

$$\beta z_i + \eta_i = E(\beta z_i + \eta_i | s_i, q_i) + (\beta \tilde{z}_i + \tilde{\eta}_i),$$

where s_i and q_i are characteristics observed by employers (i subscripts suppressed from this point on).

2. The expectation error in employers' beliefs about the worker's ability, $\beta \tilde{z} + \tilde{\eta}$, is known by the worker, but cannot be credibly signaled to prospective employers when the worker first enters the labor market.[18]

3. The public signal of the worker's ability is given by $I_t = \beta \tilde{z} + \tilde{\eta} + \nu_t$, where $\nu_t \sim N(0, \sigma_\nu^2)$. The history of public signals observed prior to production at experience level t is given by $\mathscr{F}_t = (I_0, \ldots, I_{t-1})$, which fully characterizes a worker's production history.

4. Employers use Bayesian updating based on output signals to correct their beliefs about a worker's true ability.

5. After labor market entry, workers may (in)voluntarily send a signal (d) to employers of their true ability. Employers share the prior belief that $E(\beta \tilde{z} + \tilde{\eta} | d) \neq E(\beta \tilde{z} + \tilde{\eta})$. However, as the event signal

[16] See VanderBerg (2021b) for such an application.

[17] See Appendix B.1 and Appendix B.2 for a complete treatment of this information structure and the derivations associated with it for AP's EL-SD model and the extension model developed in this paper, respectively.

[18] While not specifically a focus of this paper, there are likely numerous reasons why a worker is unable to credibly signal his/her ability a labor market entry, such as a economic/family barriers to education/job training, or a lack of incentive for high ability workers to initially pursue ways to credibly signal their ability (See e.g. Swinkels (1999) for a theoretical signaling justification behind why higher ability workers may choose to become undereducated relative to their peers). An alternate assumption similar to Frazis (2002)'s assumption that workers are initially uncertain about their ability, would be to allow the worker to imperfectly observe $\beta \tilde{z}_i + \tilde{\eta}_i$ when they enter the labor market, but to learn about it at a faster rate than firms, which would lead to the same general implications discussed below.

occurs at some experience level $t_0 > 0$, employers must update their common belief about the event signal given they have already observed $\mathscr{F}_{t_0}$. The updated belief then is given by $E[E(\beta\tilde{z}+\tilde{\eta}|d)|\mathscr{F}_{t_0}]$

Given this information structure, a worker's log-productivity at any experience level $t \geq t_0$ can be expressed as:

$$y_t = E(y_0|s,q) + \delta_d d + \tilde{f}(t) + \overbrace{E(\beta\tilde{z}+\tilde{\eta}|\mathscr{F}_{t_0})}^{\substack{\text{Expected Ability} \\ \text{Given } \mathscr{F}_{t_0}}} + \overbrace{E[E(\beta\tilde{z}+\tilde{\eta}|d)|\mathscr{F}_{t_0}]}^{\substack{\text{Expected Ability} \\ \text{Given } d \text{ at } t_0}} + \overbrace{(\beta\tilde{z}_{dt_0}+\tilde{\eta}_{dt_0})}^{\substack{\text{Remaining Error} \\ \text{Given } \mathscr{F}_{t_0} \text{ and } d}}, \tag{2.3}$$

which shows the relationship between what employers have learned about a worker's true ability prior to and at the point of an event, relative to what they still will learn as the worker's post-event experience increases.[19] Notice that after an event, as employers learn about how far away from the truth their updated beliefs are, any updating that comes from new output signals is only going to be applied to the final term related to the remaining expectation error. This implies that any effect of employer learning after t_0 on the worker's wages will come through updating this term. As such, given competition among employers, Equation (2.3) ultimately yields the log-wage process given by:

$$\omega_t = E(y_{t_0}|s,q,d_{t_0},\mathscr{F}_{t_0}) + f(t,t_0) + E(\beta\tilde{z}_d + \tilde{\eta}_d|\mathscr{F}^d_{t,t_0}) + \varepsilon_t, \tag{2.4}$$

where $f(t,t_0)$ represents the post-event experience profile of productivity for a worker who experiences the event in period t_0, and $\mathscr{F}^d_{t,t_0} = (I_{t_0},\ldots,I_{t-1})$ represents the history of output signals observed by employers since the event occurred.

2.3.3 Predictions and Implications

Define z^* as the residual from a regression of z on variables that are observed by the researcher when a worker first enters the labor market, namely s and ω_0. FG show that z^* is equivalent to observing $\tilde{z}$ for experience

[19]This equation can further be rewritten as:

$$y_t = \overbrace{E(y_{t_0}|s,q,\mathscr{F}_{t_0})}^{\substack{\text{Expected log-productivity} \\ \text{at } t_0 \text{ given } \mathscr{F}_{t_0}}} + \delta_d d + \overbrace{\tilde{f}(t,t_0)}^{\substack{\text{Post-Event} \\ \text{Exp. Profile}}} + \overbrace{E[E(\beta\tilde{z}+\tilde{\eta}|d)|\mathscr{F}_{t_0}]}^{\substack{\text{Expected Ability} \\ \text{Given } d \text{ at } t_0}} + \overbrace{(\beta\tilde{z}_{dt_0}+\tilde{\eta}_{dt_0})}^{\substack{\text{Remaining Error} \\ \text{Given } \mathscr{F}_{t_0} \text{ and } d}},$$

or as

$$y_t = \overbrace{E(y_{t_0}|s,q,d_{t_0},\mathscr{F}_{t_0})}^{\substack{\text{Expected log-productivity} \\ \text{Given } \mathscr{F}_{t_0} \text{ and } d_{t_0}}} + \overbrace{\tilde{f}(t,t_0)}^{\substack{\text{Post-Event} \\ \text{Exp. Profile}}} + \overbrace{(\beta\tilde{z}_d+\tilde{\eta}_d)}^{\substack{\text{Remaining Error} \\ \text{Given } \mathscr{F}_{t_0} \text{ and } d_{t_0}}}$$

where $\tilde{f}(t,t_0) = \tilde{f}(t) - E(\tilde{H}(t_0)|s,q,d_{t_0},\mathscr{F}_{t_0})$ represents the post-event experience profile of productivity for a worker who experiences the event in period t_0. These are a useful illustration of the fact that prior to an event signal, a worker who experiences the event will have the same learning dynamics as an identical worker who does not experience the event.

levels $t > 0$, and can be used by researchers to assess the effects of employer learning.[20,21] Consider the conditional expectation function when $t = 1, \ldots, t_0, \ldots, T$,

$$E(\omega_t | s, z^*, d_{t_0}, t) = B_{st}s + B_{zt}z^*(1 - d_{t_0}) + B'_{zt}z^*d_{t_0} + B_{dt}d_{t_0} + f(t), \tag{2.5}$$

where s, q, and z^* are reinterpreted as components of each variable that are orthogonal to $f(t)$, conditional on d_{t_0}. By separating the overall return to z^* into a non-event component and an event component, this function allows for the discrete change in the slope of the coefficient on z^* at the point of the event predicted by Equation (2.3). An alternate way of expressing the relationship between the coefficients in Equation (2.5) is by separating the CEF into components as follows,

$$\begin{aligned} E(\omega_t | s, z^*, t, d_{t_0} = 0) &= B_{st}s + B_{zt}z^* + f(t), \\ E(\omega_t | s, z^*, t, d_{t_0} = 1) &= B_{st}s + B'_{zt}z^* + B_{dt} + f(t), \end{aligned} \tag{2.6}$$

where the first equation is equivalent to the model estimated by FG with workers who experience an event excluded following the event, and the second equation is analogous to AP's EL-SD model without the restriction of starting at labor market entry.[22]

Based on these coefficients, the main predictions of the model are summarized by the following propositions.

Proposition 1 Under the assumptions of the model developed above, if an event d conveys a negative (positive) signal of a worker's ability, then at experience level t_0, (a) the value of the regression coefficient B'_{zt_0} is lower (higher) than the value of the regression coefficient B_{zt_0}, and (b) the magnitude of the difference between the two is non-increasing in t_0.

Proposition 2 Under the assumptions of the model developed above, if an event d conveys a negative (positive) signal of a worker's ability, then at all experience levels $t \geq t_0$, (a) the regression coefficient

[20] See FG or Appendix B.1 for more information.

[21] An alternate approach is taken by Light and McGee (2015b), who regress their z measures only on the s variables used in their model, leaving out the entry period wage. The advantage of that approach is that it does not require the entry period wage to be dropped in log-wage regression models based on Equation (2.5), while still purging their ability measure of any correlation to observable characteristics. However, that approach does not purge the correlation between characteristics that are only observed by the employer and the ability measure, which will complicate the interpretation of the estimated return to ability over experience, since certain aspects of learning will be correlated with these observable characteristics since their z^* is a biased measure of $\tilde{z}$. This issue is especially problematic in this context as these q measures are likely to have some form of correlation with the signaling event d, which would make it impossible to distinguish between the signaling effect of the event and this correlation.

[22] The general form of (2.5) and (2.6) may remind the reader of the structural equations used to motivate difference-in-differences models in the program evaluation literature. Indeed, the underlying intuition behind analyzing discrete changes in employer learning following the realization of some discrete signaling event is motivated in part as a way to address endogenous selection into treatment in this literature. See Abadie and Cattaneo (2018) for an excellent review of the program evaluation literature.

B'_{zt} increases at a weakly faster (slower) rate with experience than B_{zt}, and (b) the magnitude of this difference is non-increasing in t_0.

Proposition 3 Under the assumptions of the model developed above, if an event d conveys a positive signal of a worker's ability that weakly over-states $\tilde{z}$, then at all experience levels $t \geq t_0$, the regression coefficient B'_{zt_0} is non-increasing.

Proposition 4 Under the assumptions of the model developed above, if an event d conveys any signal of a worker's ability, and d is orthogonal to $f(t)$, then,

$$\frac{\partial B^d_{dt}}{\partial t} = -\delta_z \frac{\partial B'_{zt}}{\partial t} \quad \text{and} \quad \frac{\partial B^d_{dt}}{\partial t_0} = -\delta_z \frac{\partial B'_{zt}}{\partial t_0},$$

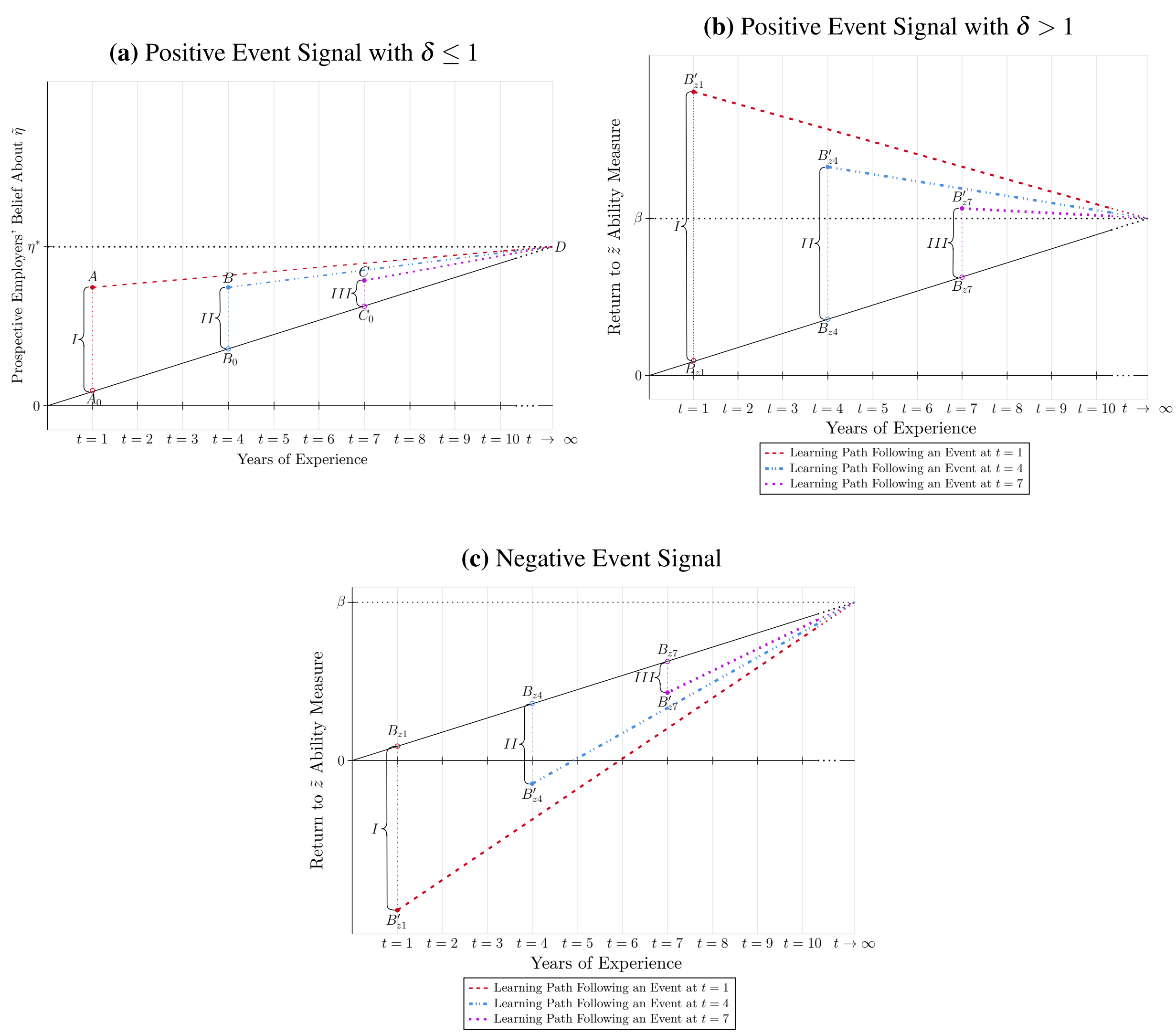
where δ_z is the regression coefficient of $\tilde{z}$ on d and can be estimated.

A formal characterization of each of the components of the regression coefficients in (2.6) and the derivations of the above model predictions can be found in Appendix B.2.

The intuition behind Proposition 1, is that when employers use events as signals of a worker's ability, they essentially are relying on a form of statistical discrimination to form their beliefs. As such, the discrete change in the return to $\tilde{\eta}$ when the event occurs is a result of employers correctly updating their beliefs about some workers, while incorrectly doing so for others. This means that workers for whom the event is a poor signal of their true ability will be disproportionately affected by the event, and thus create correlation between the residual ability measure z^* and the signaling nature of the event that is not entirely accounted for by the event indicator d in the regression equation alone. This leads directly into the intuition behind Proposition 2, which is that for those workers for whom the event was an inaccurate signal, output signals following the event should cause employers to not only place more weight on the output signal of the worker's true ability, but it should also cause them to reduce the weight placed on the signal value of the event. This is directly comparable to the intuition behind Proposition 1 in AP's statistical discrimination model. Proposition 3 can be seen as a special case of Proposition 2, whereby an event signal causes an employer to believe that a worker's ability is higher than it actually is, forcing them to then update their beliefs downward as they observe additional output signals from the worker, resulting in a decreasing return to the residual ability measure z^* over post-event experience. Figure 2.2 provides a graphical visualization of the intuition behind the predictions associated with these propositions.

The intuition above, primarily covering the (a) portions of Propositions 1 and 2, focuses on the evolution of the effect of the event at the time of the event and how it evolves with post-event experience.

Figure 2.2: Examples of Belief Paths



Notes: Each graph illustrates the effect of a different type of event signal on an employer's expectation of a worker's true unobserved ability (η^*). (a) and (b) represent positive event signals, with the signal in (a) being less than the true value of η^* while the signal in (b) overstates the true value of η^*. When the signal is positive but under states the worker's true ability, as in (a), the slope of the post-event learning path represented by the dashed red line remains positive but is smaller than it would have been in the absence of the event. In contrast, if the signal overstates ability, as in (b), the slope of the post-event learning path is negative as the employer corrects their belief downward. In (c) the event signal is negative which results in the slope of the post-event learning path to be much greater than the slope of the learning path in the absence of the event. In each of the graphs above, the magnitude of the signal decreases as potential experience increases ($I > II > III$), and the slopes of the post-event learning paths converge toward the learning path in the absence of the event.

The (b) portions of these two propositions, on the other hand, account for the fact that the longer a worker is in the labor market prior to an event, the more employers are able to update their beliefs about the worker's ability, reducing the need to rely on a (potentially inaccurate) event signal in the first place (as illustrated by differences between *I*, *II*, and *III* in Figure 2.2). The key implication here is that both pre- and post-event experience must be accounted for in order to correctly identify any event signaling that may be taking place in the labor market. This result, along with Proposition 4, which essentially describes the dynamics behind the way that employers shift weight from the signal event to the worker's actual ability, allows for direct parallels to be made between this model and AP's EL-SD model.[23]

While the symmetric employer learning assumption used for the model above provides a convenient setting for developing the key predictions of the model, it does not readily translate to situations where events are driven by the actions of an employer under asymmetric information. That said, there are a number of reasons why it may be reasonable to think that the predictions of the model above likely still provide some insight into the presence of event signaling, even though they are derived under unrealistic assumptions. First, evidence from Pinkston (2009) and L. B. Kahn (2013) suggests that learning is at least partially symmetric, and thus even if the information setting is not fully accounted for, the predictions of the model may still hold relative to the portion of learning that is symmetric in the labor market. This is increasingly relevant if employer's refuse to engage in the type of bidding wars commonly used in asymmetric learning models that lead to the lemons effect described by Gibbons and Katz (1991).[24] Additionally, evidence from Hall and Krueger (2012) suggests that only a portion of prospective employers make use of a worker's wage history during hiring, which coupled with a lack of willingness from employers to bid for their own employees, could lead wage growth to appear to be driven by symmetric employer learning for these workers, even under asymmetric information, and thus the predictions of the model above would likely hold.[25]

[23]Notice that if an event could occur at experience level $t = 0$, then Propositions 2 and 4 would be identical to AP's Propositions 1 and 2.

[24]Barron, Berger, and Black (2006) find evidence that employers are only willing to match wage offers for around 41% of their workers. Pinkston (2009) acknowledges that this issue may make the assumptions of his model (and other asymmetric employer learning models) unrealistic for all workers in the labor market. See Postel-Vinay and Robin (2004) for an example of a theoretical model that allows for some firms to bid for workers while others do not.

[25]Barach and Horton (2020) find evidence that having access to a worker's wage history directly effects how prospective employers bid for a worker. Specifically, they find that employers without access to prospective workers' wage histories evaluate applicants more intensely than employers with access to the wage histories.

2.3.4 Empirical Implementation

The basic employer learning model developed by FG and AP can be empirically estimated using the following log wage regression equation,

$$w_{iq} = \beta_0 + \beta_s s_i + \beta_{s,t}(s_i \times t_i) + \beta_z z_i^* + \beta_{z,t}(z_i^* \times t_i) + f(t_i) + \mathbf{X}_{iq} + \varepsilon_{iq}, \tag{2.7}$$

where w_{iq} is the log wage of worker i at the main employer during quarter q; s_i is years of education for worker i; t_i is a measure of total years of experience, z_i^* is a residualized ability measure, such as AFQT score; $f(t_i)$ is a cubic in experience; $\mathbf{X}$ is a vector of controls including indicators for each year-quarter pair, indicators for year interacted with education (base year of 2015), indicators for race, union status, female, entry age, and part-time status, as well as indicators for entry quarter two-digit industry.

The key intuition behind this regression is that if the employer learning models are correct, then if z_i^* is a residualized ability measure, we would expected β_z to be zero and $\beta_{z,t} > 0$ as the effect of the unobserved component of ability is learned by the labor market, and is thus factored into the wage equation. Additionally, as the ability measure is orthogonal to information available to the employers when the worker first enters the labor market, $\beta_{s,t}$ should be equal to zero as the labor market returns to education should be fully captured in the first period as it is fully observed by the market.[26]

The employer learning with event signals model developed in Section 2.3 is more nuanced. The model can be empirically estimated using the following log wage regression equation,

$$\begin{aligned}\omega_{iq} = \; & \beta_0 + \beta_s s_i + \beta_z z_i^* + \beta_{s,t}(s_i \times t_i) + \beta_{z,t}(z_i^* \times t_i) + \beta_d d_{it_0} + \beta_{z,d}(z_i^* \times d_{it_0}) \\ & + \beta_{z,t'}(z_i^* \times d_{it_0} \times t_i') + \beta_{z,t_0}(z_i^* \times d_{it_0} \times t_0) + \beta_{z,t_0,t'}(z_i^* \times d_{it_0} \times t_0 \times t_i') \\ & + f(t_i) + \tilde{f}(t_0, t_i') + \mathbf{X}_{iq} + \varepsilon_{iq},\end{aligned} \tag{2.8}$$

where $d_{it_0} = \mathbb{1}(t \geq t_0)$ is an indicator for all periods following some labor market event, t_0 represents years of pre-event experience, t_i' represents post-event experience for worker i, and $\tilde{f}(t_0, t_i')$ is a function of pre- and post-event experience. To match the asymmetric employer learning models of Schönberg (2007a) or

[26] Technically these implications only hold for the FG model. In the EL-SD model developed in AP, where z_i is used in place of z_i^*, the expected coefficients have a slightly different interpretation, and this interpretation will be useful when interpreting the estimates from the asymmetric learning model developed in Section 2.3. In the EL-SD model, β_3 need not equal zero as some portion of ability related to, e.g. education, is observed by the market. Given that education, and the ability measure are likely correlated in the model, if employers have an underlying belief about a worker's ability, conditional on some observed variable (education), we should find $\beta_4 > 0$, while $\beta_2 < 0$ as the market shifts the weight it puts on the relationship between education and ability when a worker first enters the market to the observed productivity signals for the worker overtime. That is, as a worker gains more experience, the market should have a better understanding of her innate ability, and will no longer need to rely on its beliefs about ability, given observed education.

Pinkston (2009), this equation can be modified to include the interaction of z^* and x_c, and/or the interaction of z^* and j_c^s, where j_c^s represents the length of a worker's current job spell.

The key aspects of this regression model are in how it relates to the combined return to ability following an event (Appendix Equation (B.19)) and the results from Propositions 1 - 4. In this regression model, β_{zt} represents the experience-ability profile for workers for whom $d_{it_0} = 0$ ($\frac{\partial B_{zt}}{\partial t}$); $\beta_{z,d}$ represents the experience-invariant event signal (b_{zd}); $\beta_{z,t'}$ represents the return to the post-event experience-ability profile ($\frac{\partial B'_{zt}}{\partial t'}$); β_{z,t_0} represents the return to the pre-event experience-ability ($\frac{\partial B'_{zt}}{\partial t_0}$); and $\beta_{z,t_0,t'}$ represents the way in which pre- and post-event experience interact with each other and the ability measure ($\frac{\partial^2 B'_{zt}}{\partial t' \partial t_0}$). If an event is such that $\delta_z < 0$, the model predicts that $\beta_{z,d} < 0$, while $\beta_{z,t'}, \beta_{z,t_0} > \beta_{zt} > 0$, and $\beta_{z,t_0,t'} < 0$. Conversely if $\delta_z > 1$, the model predicts that $\beta_{z,d} > 0$, while $\beta_{z,t'}, \beta_{z,t_0} < 0 < \beta_{zt}$, and $\beta_{z,t_0,t'} > 0$.

2.4 Data

The data used for this analysis comes from the 2015 release of the National Longitudinal Survey of Youth 1997 cohort (NLSY97), a nationally representative survey of 8,984 men and women who were between the ages of 12 and 16 on December 31st, 1996. This survey, which was designed to capture the evolution of employment career paths of individuals from the time they leave school through adulthood, was administered annually between 1997 and 2011 before switching to its current biennial format following the completion of the 2011 interview round. During these interviews, extensive events histories are collected from the respondents related to a variety of topics covering employment, program participation, and education, as well as other important life events such as marital status and parental cohabitation. The event history data available in the NLSY97, as well as the detailed employer roster,[27] make this an ideal source from which to study the effects of mid-career education changes.

The key feature of the NLSY97, relative to other potential data sources, is the information available on pre-market skills, aptitude and cognition tests (such as the ASVAB), as well detailed histories on a variety of topics such as incarceration or drug use, which can be used as proxies for ability/quality in this analysis. AFQT scores (which are derived from components of the ASVAB) are used by both FG and AP (among others) as variables which are correlated with ability, but are likely unobserved by employers at labor market entry. See Lange (2007) for a discussion of the usefulness of using AFQT scores as measures of ability that

[27] The employer roster provides information on a variety of employment characteristics for each job an individual works ranging from the industry/occupation to whether the worker enjoys his/her job. Addition-ally, each employer on the roster is assigned a unique identifier which can be matched with the employer identifiers used in the event history data.

are unobserved by employers.

In order to analyze the relationship between employer learning and increased educational attainment through wage changes, I restrict my sample to individuals who have made their first long-term transition from school to the labor market prior to 2011.[28] Actual experience is defined as the number of weeks an individual has worked at least 30 hours divided by 50, while my primary measure of potential experience is defined as the number of quarters since an individual began their first employment spell divided by four.

For cognitive ability variables, I use both the ASVAB percentile score computed directly in the NLSY97 data, as well as the adjusted AFQT scores created by Altonji et al. (2012a) that is directly comparable to the AFQT measure used in FG and AP, which is a standard measure of a worker's productive ability used in the literature dating back to Neal and Johnson (1996).[29] Both of these measures are age adjusted and normalized to have mean zero and unit variance in the overall sample.

To create measures of these ability variables that are orthogonal to the information available to the prospective employers when a worker first enters the labor market, I create a sample consisting of each individual's first period in the labor market (i.e. the first period they are employed and report a non-zero wage). Using this sample, I compute residual scores for the ability variables by regressing each measure on a vector of observable characteristics which should be easily seen by the market during the period, as well as the log wage earned by the individual that quarter.[30,31] This regression accounts for approximately 40 percent of the variation in AFQT scores, which is noticeably lower than the R^2 value found in FG (53 percent of variation accounted for using their NLSY79 sample and similar controls), but is in line with recent empirical evidence that suggests that the return to cognitive ability has generally been decreasing over the past few decades (see e.g. Castex and Dechter (2014), and Beaudry et al. (2016)). I will show later, however, that this finding (as well as the apparent decline in the returns to cognitive ability found previously) may be driven in part by an apparent lack of updating regarding a worker's unobserved ability from the perspective of the

[28]I Define this transition as being the first quarter in which an individual does not increase their education level the following year and will have worked at least 30 hours per week for half of the weeks during the following two years.

[29]There are some differences between the two measures as the number of respondents who completed the required sections for the AFQT scores is less than for the ASVAB percentile score.

[30]The vector of observable characteristics contains four education dummy variables ($<$ 12 years, 12 years, 13-15 years, and 16 years), an indicator for part-time status, the interaction of part-time status and each education dummy, indicators for race, sex, age ($<$18, 18-21, 22-26, $>$26), year, quarter, each year/quarter interaction, two-digit industry, the log of the worker's wage as well as each of these interacted with the education dummy variables. Additionally, I include the log number of employees at the entry period employer as a number of recent studies have found this to be an important determinant of labor market career trajectories. See e.g. Moscarini and Postel-Vinay (2012) and Arellano-Bover (2020), though Haltiwanger et al. (2018) find minimal evidence of firm size job ladders.

[31]FG use the level of a worker's wage in their analysis instead of using its log. As my theoretical model from section 2.3 follows AP by using a worker's log wage, my main analysis will as well.

labor market during the first half decade of his/her career.

The return to school sample also consists of 4,576 individuals divided into subsamples based on the type of degree they obtained (if any) after making an initial long-term transition into the labor market.[32]. Among this sample, 859 respondents increase their education level after entering the sample and report attaining a new degree level (33,745 respondent/quarter obs). Table 2.1 reports summary statistics for this group during their first quarter in the sample, broken down by the type of degree attained. While there are mechanical differences in many of the entry quarter variables related to different levels of entry education levels (e.g. worker's who return to school to earn there GED mechanically have less than 12 years of education), the mean residual AFQT scores are quite intriguing. Among the entire return to school sample, the mean residual AFQT score is substantially higher than for the control, ranging from 0.06 standard deviations higher for the GED sample to around 0.32 standard deviations higher for the BA Degree sample. Table 2.2 reports summary statistics for several of these variables over the entire sample period.

2.5 Returning to School

In attempting to identify the signaling role of returning to education, a more nuanced approach is needed than in the previous subsection due to the fact that the signal component ($\delta_z^{RS} > 0$) has the same sign as the true return to ability (β), which means that if the signaling role of returning to school is rather minor, the components of the pre- and post-schooling experience-ability profiles will cancel each other out leading to the appearance of no effect. Basically, what this means is that if the signal is weak, the pre- and post-schooling experience-ability profiles will be closer to zero than the overall experience-ability profile, which will make point identification challenging, if not impossible.[33] While the estimated interaction between ability and an indicator for returning to school may still be identified in the case of a weak signal, distinguishing the effect of the signal from differential effects of human capital accumulation based on residual ability is not possible.

While identification is an issue if the signal is weak, it is not if the signal is sufficiently strong ($\delta_z^{RS} > 1$).

[32]Note that approximately 10,000 respondent/quarter observations are dropped for individuals who change education levels but do not report receiving a new degree following entry into the sample

[33]One potential work around to this problem is to consider the total experience-ability profiles for workers who do return to school versus those who do not. While not a formal prediction of the model due to the discontinuous derivative of the return to ability with respect to total experience for workers who return to school around the point of the event, a regression that compares the total experience-ability profiles of the two types of worker should find that workers who return to school have a much larger estimated ability-experience interaction as the correlation between ability and returning to school is positive, and the correlation between experience and returning to school is also positive, which will increase the estimated effect for workers who returned to school, while decreasing the estimated effect for workers who did not.

Table 2.1: Entry Quarter Summary Statistics by Return to Educ Sample

	Control	GED/HS	AA/BA	Grad
Wage	12.78	9.533	11.14	16.10
	[7.444]	[2.979]	[4.337]	[6.787]
Log Wage	2.444	2.216	2.343	2.689
	[0.424]	[0.274]	[0.365]	[0.438]
Years of Education	13.63	10.14	13.35	16.34
	[2.365]	[0.922]	[1.435]	[0.934]
Age	21.02	18.54	20.01	22.73
	[2.709]	[2.101]	[2.049]	[1.565]
Female	0.485	0.418	0.620	0.652
	[0.500]	[0.495]	[0.486]	[0.477]
Black	0.244	0.283	0.270	0.210
	[0.430]	[0.451]	[0.444]	[0.408]
Hispanic	0.204	0.304	0.176	0.129
	[0.403]	[0.461]	[0.381]	[0.336]
Part-Time	0.180	0.196	0.251	0.155
	[0.384]	[0.398]	[0.434]	[0.362]
Union	0.0961	0.0707	0.0880	0.167
	[0.295]	[0.257]	[0.284]	[0.374]
Urban	0.793	0.786	0.770	0.839
	[0.405]	[0.412]	[0.421]	[0.368]
Size of Employer	346.0	185.9	181.9	573.9
	[1684.7]	[668.6]	[494.3]	[1662.4]
Standardized AFQT Score	-0.0463	-0.718	0.229	0.768
	[1.009]	[0.934]	[0.828]	[0.621]
Std. Residual AFQT Score	-0.0508	0.115	0.246	0.141
	[1.007]	[1.047]	[0.989]	[0.739]
Observations	3654	184	534	233

Note: Standard deviations in parentheses.
Source: Author's tabulations of NLSY97 data. See Section 2.4 for information regarding the construction of the sample presented here.

Thus, with a sufficiently strong signal, Propositions 1-4 provide a number of predictions which can be used to identifying the signaling nature of returning to school. First, the coefficient on the interaction between an ability measure and the return to school indicator should be positive. This should be rather intuitive as the cost of returning to school should deter individuals with lower ability from pursuing additional education, thus this should disproportionately effect workers with high residual ability. Second, the coefficients on the interactions between both pre- and post-schooling experience interacted with ability should be negative ($\frac{\partial B^{RS}_{zt}}{\partial t_0}, \frac{\partial B^{RS}_{zt}}{\partial t'} < 0$). While this may seem surprising, consider that a strong signal will cause some employers' expectations to essentially over-shoot their worker's true ability, which means that they will be updating their

Table 2.2: All Year-Quarter Summary Statistics by Return to Educ Sample

	Control	GED/HS	AA/BA	Grad
Wage	17.58	13.48	17.11	25.02
	[10.54]	[6.783]	[9.186]	[13.10]
Log Wage	2.738	2.510	2.724	3.104
	[0.486]	[0.409]	[0.470]	[0.476]
Potential Experience	6.293	7.803	7.261	6.141
	[4.426]	[5.161]	[4.744]	[3.955]
Actual Experience	5.416	5.484	4.403	3.945
	[3.931]	[4.173]	[3.382]	[2.856]
Entry Education	13.55	10.17	13.30	16.27
	[2.320]	[0.937]	[1.389]	[0.930]
Years of Education	13.55	11.42	14.79	17.64
	[2.320]	[1.384]	[1.954]	[1.413]
Current Job Spell Length	3.006	2.432	2.903	3.455
	[3.172]	[3.067]	[3.075]	[3.147]
Tenure With Current Employer	2.602	2.192	2.359	3.018
	[2.864]	[2.696]	[2.618]	[2.931]
Age	26.97	25.93	27.06	28.64
	[4.591]	[5.145]	[4.801]	[3.962]
Standardized AFQT Score	-0.0517	-0.687	0.247	0.768
	[1.004]	[0.907]	[0.827]	[0.610]
Std. Residual AFQT Score	-0.0730	0.0821	0.256	0.144
	[1.013]	[1.015]	[0.999]	[0.729]
Unique Individuals	3,654	184	534	233
Observations	137,228	7,911	23,720	9,670

Note: Standard deviations in parentheses.
Source: Author's tabulations of NLSY97 data. See Section 2.4 for information regarding the construction of the sample presented here.

beliefs downward as the worker's true ability is gradually revealed with pre- and/or post-event experience. Finally, the third prediction is that the coefficient on the interaction between pre-$\times$post-schooling experience and ability should be positive ($\frac{\partial^2 B^{RS}_{zt}}{\partial t_0 \partial t'} > 0$).

2.5.1 Estimating The Signaling Role of Returning to School

Table 2.3 reports estimates based on the predictions regarding the signaling role of returning to school discussed in Section 2.5, broken down into three subgroups based on the degree that the worker attained after returning the school. The columns in Table 2.3 gradually add to the ability-experience profile for workers who return to school, with Column 7 reporting the estimates of the regression model in Equation (2.8), and Column 8 replicating these results using wage levels instead of log wage. Note that the pre- and post-return to school experience-AFQT interaction terms are the per year estimates, not the 10 year estimates used in the traditional employer learning literature (*e.g.* Altonji and Pierret 2001).

For workers who return to school to get their GED, the estimates in Column 8 match the predictions of from Section 2.5 regarding the signaling nature of returning to school. The coefficient on the AFQT-event interaction term is positive (0.299 log points) and significant (at the 0.1% level), while the pre-GED experience-AFQT interaction is negative (-0.04) and significant at the 1% level. The other coefficients corresponding to the predictions of the model all have the correct sign, however they are imprecisely estimated.

For workers who return to school to get their Associates or Bachelors degrees, the estimates in Table 2.3 provide no evidence that signaling is taking place. For the main specifications in Column 8, the coefficient on the AFQT-event interaction term for both events is negative and imprecisely estimated. This runs counter to the prediction that the coefficient should be positive if signaling is taking place. Further, the signs of the coefficients on the ability-experience parameters do not match the predictions from the model.

For workers who return to school to get a graduate degree, the estimates in Column 8 match the predictions of the model above. The coefficient on the AFQT-event interaction is positive (0.220 log points) and significant at the 5% level, and this effect is decreasing in pre-event experience (-0.045 log points at the 5% significance level). Following the event, the AFQT post-event experience interaction is negative (-0.006 log points) but imprecisely estimated, while the coefficient on the AFQT pre$\times$post-event experience variable being positive (0.0007) but also imprecisely estimated. Each of these results match the predictions discussed above and provide suggestive evidence in support of the signaling role of returning to school for a graduate degree.

From a practical standpoint, these results indicate that a standard deviation increase in residual AFQT score among workers who return to school to get their GED or graduate degree experience leads to between 19 percent (GED) and 28 percent (graduate) higher wages after the event, with this effect decreasing by between three (GED) and six percentage points per year of pre-event experience. Following the event, a standard deviation increase in residual AFQT score among these workers leads to a wage decrease of between one (GED) and four (graduate) percent each year following the event, with this effect decreasing toward zero by two percentage points per year of pre-event experience for the workers who get their graduate degree. These results match up well with the findings of Tyler et al. (2000) and Hussey (2012) who find that

Table 2.3: Return to School Estimation — Log Wage Regressions Using Potential Experience

Independent Variable	(1) Log Wage	(2) Log Wage	(3) Log Wage	(4) Log Wage	(5) Log Wage	(6) Log Wage	(7) Log Wage	(8) Wage
Education	0.091***	0.092***	0.092***	0.093***	0.093***	0.091***	0.091***	2.022***
	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.396)
Educ × Total Exp/10	0.013	0.009	0.009	0.008	0.008	0.010	0.010	0.546^{+}
	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.011)	(0.011)	(0.302)
AFQT*	0.001	0.002	0.003	0.001	0.002	0.003	0.003	-0.148
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.143)
AFQT* × Total Exp/10	0.041***	0.038***	0.036***	0.040***	0.040***	0.038***	0.038***	1.171***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)	(0.247)
GED/HS Deg		0.088	-0.003	-0.033	-0.073	-0.528**	-0.423*	-8.790**
		(0.070)	(0.070)	(0.070)	(0.085)	(0.166)	(0.209)	(3.175)
GED/HS Deg × Post Exp			0.019	0.018	0.025	0.027	0.009	0.187
			(0.015)	(0.015)	(0.019)	(0.017)	(0.037)	(0.669)
GED/HS Deg × AFQT*				0.075^{+}	0.099*	0.251**	0.299***	5.164**
				(0.043)	(0.050)	(0.087)	(0.089)	(1.784)
GED/HS Deg × AFQT* × Post Exp					-0.005	-0.005	-0.018	-0.371^{+}
					(0.008)	(0.008)	(0.014)	(0.206)
GED/HS Deg × AFQT* × Pre Exp						-0.026^{+}	-0.040**	-0.826**
						(0.013)	(0.014)	(0.251)
GED/HS Deg × AFQT* × Pre × Post Exp							0.004	0.066
							(0.003)	(0.040)
AA/BA Deg		0.083^{+}	0.007	0.017	-0.030	-0.067	0.139	2.964
		(0.047)	(0.046)	(0.046)	(0.048)	(0.090)	(0.099)	(2.182)
AA/BA Deg × Post Exp			0.028**	0.027**	0.036***	0.032**	-0.018	-0.481
			(0.009)	(0.009)	(0.011)	(0.011)	(0.019)	(0.476)
AA/BA Deg × AFQT*				-0.052^{+}	-0.043^{+}	-0.107^{+}	-0.083	-2.169^{+}
				(0.027)	(0.026)	(0.056)	(0.052)	(1.177)
AA/BA Deg × AFQT* × Post Exp					-0.002	0.005	-0.000	-0.202
					(0.005)	(0.004)	(0.008)	(0.222)
AA/BA Deg × AFQT* × Pre Exp						0.014^{+}	0.009	0.194
						(0.008)	(0.008)	(0.178)
AA/BA Deg × AFQT* × Pre × Post Exp							0.001	0.034
							(0.002)	(0.047)
Grad Deg		0.082	0.107^{+}	0.104^{+}	0.096	-0.082	-0.098	-2.018
		(0.060)	(0.059)	(0.059)	(0.061)	(0.258)	(0.268)	(7.503)
Grad Deg × Post Exp			-0.018	-0.019	-0.016	-0.015	-0.011	0.419
			(0.016)	(0.016)	(0.017)	(0.017)	(0.031)	(1.006)
Grad Deg × AFQT*				0.059^{+}	-0.021	0.117	0.220*	4.904
				(0.031)	(0.038)	(0.090)	(0.107)	(3.084)
Grad Deg × AFQT* × Post Exp					0.022**	0.021**	-0.006	0.354
					(0.008)	(0.007)	(0.018)	(0.606)
Grad Deg × AFQT* × Pre Exp						-0.022	-0.045*	-1.018
						(0.017)	(0.021)	(0.651)
Grad Deg × AFQT* × Pre × Post Exp							0.007	0.075
							(0.004)	(0.147)
R^2	0.441	0.445	0.446	0.447	0.448	0.449	0.450	0.407
Observations	114,167	114,167	114,167	114,167	114,167	114,167	114,167	114,167
Individuals	4,346	4,346	4,346	4,346	4,346	4,346	4,346	4,346
No. of GED/HS Degrees	.	100	100	100	100	100	100	100
No. of AA/BA Degrees	.	340	340	340	340	340	340	340
No. of Graduate Degrees	.	166	166	166	166	166	166	166

Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors are clustered at the individual worker level. The first year of potential experience is dropped from the analysis as this was used to create the residual ability measures. All models include a quadratic in pre- and post- return to school potential experience, indicators for each type of degree received, pre-return to school experience interacted with post-return to school experience, a vector of year-quarter indicators (base 2015), education interacted with a vector of year indicators, indicators for race, female, union status, part-time status, two-digit entry industry, and the log number of employees at the workers entry job, all interacted with a cubic time trend (base 2015). With the exception of the two-digit entry industry dummies, each individual control above is also interacted with indicators for each type of return to school.

ability signaling plays a significant role in the overall returns to obtaining a GED and an MBA, respectively.

2.5.2 Robustness Checks

Table 2.4 provides estimation results based on Equation (2.8) for without including controls for the interaction between observable characteristics and the event experience profile. As with the main results reported above, the estimates in Column 8 match the predictions of the model for workers who return to school to get a GED or graduate degree, but not for those who return to school for an Associates or Bachelors degree. The estimates from this model are generally stronger than those from the main model, with the pre and pre-post-event ability interactions significantly estimated at the 10 % level or better for the GED sample and at the 5% level or better for the graduate degree sample. While my preferred specification allows for different event experience paths based on observable characteristics, the results presented here for the model without these additional controls provide further evidence in support of the results found above for the signaling nature of of returning to school to get a GED or graduate degree.

Tables 2.5 and 2.6 back up the results of Tables 2.3 and 2.4 by replicating there results using actual experience instrumented by potential experience. The IV estimates are generally as strong, if not stronger than the results of the primary specification.

2.6 Conclusion

In this paper, I extend the traditional employer learning models of H. Farber and Gibbons (1996) and Altonji and Pierret (2001) to allow post-labor market entry events to serve as worker ability signals. The model allows for the identification of post-labor market entry statistical discrimination as employers update prior beliefs about a given worker's ability.

Consistent with Tyler et al. (2000) and Hussey (2012), results from an empirical application suggest that returning to school to receive a GED or graduate degree sends a positive ability signal to prospective employers. I contribute to a limited amount of existing research on the signaling role of returning to school (Tyler et al. 2000; Hussey 2012). Given the large proportion of workers who return to school after entering the labor market (Light 1995; Light and McGee 2015a), future research that seeks to better understand how mid-career signaling events impact wage and earnings trajectories is warranted.

In future work, researchers may use the approach I develop in this paper to examine the extent to which signaling effects differ across workers with different skills and demographic characteristics, which have been

Table 2.4: Return to School Estimation - No Control Interactions — Log Wage Regressions Using Potential Experience

Independent Variable	(1) Log Wage	(2) Log Wage	(3) Log Wage	(4) Log Wage	(5) Log Wage	(6) Log Wage	(7) Log Wage	(8) Wage
Education	0.091^{***}	0.095^{***}	0.095^{***}	0.096^{***}	0.096^{***}	0.095^{***}	0.094^{***}	2.095^{***}
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.390)
Educ × Total Exp/10	0.013	0.008	0.007	0.006	0.006	0.007	0.008	0.512^{+}
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.298)
AFQT*	0.001	0.002	0.003	0.002	0.001	0.003	0.003	-0.144
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.143)
AFQT* × Total Exp/10	0.041^{***}	0.038^{***}	0.036^{***}	0.040^{***}	0.041^{***}	0.038^{***}	0.038^{***}	1.165^{***}
	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)	(0.247)
GED/HS Deg		0.068^{+}	-0.015	-0.019	-0.023	-0.032	-0.015	-1.013
		(0.040)	(0.042)	(0.040)	(0.039)	(0.038)	(0.112)	(1.683)
GED/HS Deg × Post Exp			0.022	0.022	0.023	0.022	-0.010	-0.119
			(0.015)	(0.015)	(0.015)	(0.015)	(0.020)	(0.344)
GED/HS Deg × AFQT*				0.052	0.091^{+}	0.166^{+}	0.215^{*}	3.644^{+}
				(0.039)	(0.051)	(0.087)	(0.096)	(1.886)
GED/HS Deg × AFQT* × Post Exp					-0.008	-0.006	-0.018	-0.459^{*}
					(0.006)	(0.006)	(0.011)	(0.192)
GED/HS Deg × AFQT* × Pre Exp						-0.013	-0.029	-0.575^{+}
						(0.013)	(0.018)	(0.314)
GED/HS Deg × AFQT* × Pre × Post Exp							0.004^{+}	0.091^{*}
							(0.003)	(0.042)
AA/BA Deg		0.059^{*}	-0.022	-0.009	-0.012	-0.010	0.079	2.307
		(0.023)	(0.026)	(0.026)	(0.025)	(0.025)	(0.074)	(1.569)
AA/BA Deg × Post Exp			0.027^{**}	0.026^{**}	0.027^{**}	0.026^{**}	-0.010	-0.423
			(0.009)	(0.009)	(0.009)	(0.009)	(0.015)	(0.360)
AA/BA Deg × AFQT*				-0.052^{+}	-0.041	-0.107^{+}	-0.080	-2.068
				(0.027)	(0.026)	(0.057)	(0.055)	(1.289)
AA/BA Deg × AFQT* × Post Exp					-0.003	0.005	-0.002	-0.240
					(0.005)	(0.004)	(0.008)	(0.229)
AA/BA Deg × AFQT* × Pre Exp						0.014^{+}	0.008	0.180
						(0.009)	(0.009)	(0.203)
AA/BA Deg × AFQT* × Pre × Post Exp							0.002	0.043
							(0.002)	(0.049)
Grad Deg		0.102^{***}	0.124^{***}	0.115^{***}	0.126^{***}	0.129^{***}	0.020	-1.548
		(0.030)	(0.033)	(0.034)	(0.033)	(0.033)	(0.187)	(5.330)
Grad Deg × Post Exp			-0.020	-0.020	-0.022	-0.022	-0.017	-0.073
			(0.017)	(0.016)	(0.016)	(0.016)	(0.025)	(0.833)
Grad Deg × AFQT*				0.076^{*}	-0.014	0.131	0.261^{**}	6.183^{*}
				(0.030)	(0.038)	(0.087)	(0.100)	(2.873)
Grad Deg × AFQT* × Post Exp					0.024^{**}	0.024^{**}	-0.010	-0.060
					(0.008)	(0.008)	(0.017)	(0.610)
Grad Deg × AFQT* × Pre Exp						-0.023	-0.052^{**}	-1.251^{*}
						(0.016)	(0.020)	(0.613)
Grad Deg × AFQT* × Pre × Post Exp							0.008^{*}	0.172
							(0.004)	(0.145)
R^2	0.441	0.443	0.444	0.445	0.446	0.446	0.447	0.403
Observations	114,167	114,167	114,167	114,167	114,167	114,167	114,167	114,167
Individuals	4,346	4,346	4,346	4,346	4,346	4,346	4,346	4,346
No. of GED/HS Degrees	.	100	100	100	100	100	100	100
No. of AA/BA Degrees	.	340	340	340	340	340	340	340
No. of Graduate Degrees	.	166	166	166	166	166	166	166

Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors are clustered at the individual worker level. The first year of potential experience is dropped from the analysis as this was used to create the residual ability measures. All models include a quadratic in pre- and post- return to school potential experience, indicators for each type of degree received, pre-return to school experience interacted with post-return to school experience, a vector of year-quarter indicators (base 2015), education interacted with a vector of year indicators, indicators for race, female, union status, part-time status, two-digit entry industry, and the log number of employees at the workers entry job, all interacted with a cubic time trend (base 2015).

shown to affect employer learning.[34] Heterogeneous effects by worker characteristics could prove important if, for instance, there is a higher return to signaling certain types of skills that are difficult for employers to observe. Moreover, understanding how signaling effects vary by race and gender may inform policies intended to reduce existing wage gaps. Hence, a more thorough understanding of the role of mid-career signaling will help to inform workers about expected wage returns to returning to school and other activities that send positive ability signals to prospective employers.

[34] See Light and McGee (2015b) for evidence on differences in the way employers learn about different skills; and Pinkston (2003) and Arcidiacono et al. (2010) for differences in learning across demographic groups.

Table 2.5: Return to School Estimation — Log Wage Regressions Using Actual Experience Instrumented with Potential Experience

Independent Variable	(1) Log Wage	(2) Log Wage	(3) Log Wage	(4) Log Wage	(5) Log Wage	(6) Log Wage	(7) Log Wage	(8) Wage
Education	0.074^{***}	0.092^{***}	0.091^{***}	0.092^{***}	0.092^{***}	0.089^{***}	0.089^{***}	1.981^{***}
	(0.015)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.414)
Educ × Total Exp/10	0.021	0.000	0.002	0.001	0.001	0.004	0.004	0.461
	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.354)
AFQT*	0.002	0.004	0.005	0.004	0.004	0.005	0.005	0.083
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.124)
AFQT* × Total Exp/10	0.059^{***}	0.044^{***}	0.043^{***}	0.044^{***}	0.044^{***}	0.041^{***}	0.041^{***}	1.021^{***}
	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.271)
GED/HS Deg		0.135^{+}	0.040	0.006	-0.004	-0.311^{*}	-0.187	-4.143
		(0.070)	(0.071)	(0.069)	(0.073)	(0.146)	(0.170)	(2.551)
GED/HS Deg × Post Exp			0.030	0.029	0.034	0.037	-0.005	0.019
			(0.020)	(0.020)	(0.025)	(0.024)	(0.046)	(0.774)
GED/HS Deg × AFQT*				0.082^{+}	0.116^{*}	0.257^{**}	0.270^{***}	4.700^{**}
				(0.042)	(0.047)	(0.080)	(0.078)	(1.494)
GED/HS Deg × AFQT* × Post Exp					-0.013	-0.016	-0.024	-0.363
					(0.013)	(0.012)	(0.021)	(0.328)
GED/HS Deg × AFQT* × Pre Exp						-0.031^{+}	-0.039^{**}	-0.717^{**}
						(0.018)	(0.015)	(0.249)
GED/HS Deg × AFQT* × Pre × Post Exp							0.004	0.050
							(0.006)	(0.115)
AA/BA Deg		0.217^{***}	0.183^{***}	0.190^{***}	0.153^{**}	0.135^{*}	0.210^{**}	3.571^{*}
		(0.048)	(0.048)	(0.048)	(0.049)	(0.066)	(0.068)	(1.465)
AA/BA Deg × Post Exp			0.024^{*}	0.023^{*}	0.032^{**}	0.024^{*}	-0.003	-0.076
			(0.010)	(0.010)	(0.012)	(0.012)	(0.017)	(0.411)
AA/BA Deg × AFQT*				-0.034	-0.026	-0.077^{+}	-0.066^{+}	-1.794^{*}
				(0.027)	(0.026)	(0.044)	(0.040)	(0.902)
AA/BA Deg × AFQT* × Post Exp					-0.003	0.006	0.002	-0.044
					(0.005)	(0.005)	(0.007)	(0.186)
AA/BA Deg × AFQT* × Pre Exp						0.019	0.015	0.508^{+}
						(0.014)	(0.012)	(0.282)
AA/BA Deg × AFQT* × Pre × Post Exp							0.002	0.061
							(0.004)	(0.111)
Grad Deg		0.250^{***}	0.279^{***}	0.273^{***}	0.264^{***}	0.180	0.135	2.761
		(0.061)	(0.059)	(0.059)	(0.059)	(0.163)	(0.168)	(4.379)
Grad Deg × Post Exp			-0.024	-0.024	-0.019	-0.023	-0.009	0.483
			(0.016)	(0.016)	(0.017)	(0.017)	(0.022)	(0.740)
Grad Deg × AFQT*				0.081^{**}	0.009	0.113^{+}	0.163^{**}	3.863^{*}
				(0.031)	(0.037)	(0.060)	(0.061)	(1.697)
Grad Deg × AFQT* × Post Exp					0.025^{**}	0.025^{**}	0.008	0.646^{+}
					(0.009)	(0.008)	(0.011)	(0.360)
Grad Deg × AFQT* × Pre Exp						-0.038^{+}	-0.065^{**}	-1.349^{+}
						(0.023)	(0.025)	(0.733)
Grad Deg × AFQT* × Pre × Post Exp							0.010	0.156
							(0.007)	(0.224)
R^2	0.430	0.455	0.455	0.457	0.457	0.456	0.456	0.407
Observations	114,167	114,167	114,167	114,167	114,167	114,167	114,167	114,167
Individuals	4,346	4,346	4,346	4,346	4,346	4,346	4,346	4,346
No. of GED/HS Degrees	.	100	100	100	100	100	100	100
No. of AB Degrees	.	340	340	340	340	340	340	340
No. of Graduate Degrees	.	166	166	166	166	166	166	166

Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.6: Return to School Estimation - No Control Interactions — Log Wage Regressions Using Actual Experience Instrumented with Potential Experience

Independent Variable	(1) Log Wage	(2) Log Wage	(3) Log Wage	(4) Log Wage	(5) Log Wage	(6) Log Wage	(7) Log Wage	(8) Wage
Education	0.074***	0.095***	0.094***	0.095***	0.094***	0.093***	0.093***	2.071***
	(0.015)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.406)
Educ × Total Exp/10	0.021	-0.002	-0.000	-0.001	-0.001	0.000	0.001	0.413
	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.348)
AFQT*	0.002	0.004	0.005	0.004	0.003	0.005	0.005	0.085
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.124)
AFQT* × Total Exp/10	0.059***	0.044***	0.042***	0.043***	0.044***	0.041***	0.040***	1.019***
	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.271)
GED/HS Deg		0.126**	0.049	0.045	0.040	-0.041	0.028	-0.002
		(0.040)	(0.039)	(0.037)	(0.035)	(0.097)	(0.079)	(1.276)
GED/HS Deg × Post Exp			0.034^{+}	0.034^{+}	0.037^{+}	0.035^{+}	0.009	0.266
			(0.020)	(0.020)	(0.019)	(0.019)	(0.023)	(0.384)
GED/HS Deg × AFQT*				0.055	0.105*	0.188*	0.213**	3.630*
				(0.040)	(0.048)	(0.080)	(0.081)	(1.632)
GED/HS Deg × AFQT* × Post Exp					-0.017^{+}	-0.019^{+}	-0.029^{+}	-0.550^{+}
					(0.010)	(0.010)	(0.016)	(0.296)
GED/HS Deg × AFQT* × Pre Exp						-0.017	-0.030^{+}	-0.502
						(0.016)	(0.017)	(0.320)
GED/HS Deg × AFQT* × Pre × Post Exp							0.006	0.125
							(0.005)	(0.099)
AA/BA Deg		0.207***	0.167***	0.175***	0.173***	0.126**	0.174***	3.304***
		(0.024)	(0.025)	(0.025)	(0.025)	(0.043)	(0.042)	(0.894)
AA/BA Deg × Post Exp			0.024*	0.024*	0.024*	0.018^{+}	-0.003	-0.158
			(0.010)	(0.010)	(0.010)	(0.010)	(0.014)	(0.338)
AA/BA Deg × AFQT*				-0.035	-0.024	-0.076^{+}	-0.066	-1.785^{+}
				(0.027)	(0.026)	(0.044)	(0.041)	(0.952)
AA/BA Deg × AFQT* × Post Exp					-0.003	0.005	0.002	-0.047
					(0.006)	(0.005)	(0.007)	(0.193)
AA/BA Deg × AFQT* × Pre Exp						0.020	0.015	0.493
						(0.014)	(0.013)	(0.308)
AA/BA Deg × AFQT* × Pre × Post Exp							0.003	0.086
							(0.004)	(0.113)
Grad Deg		0.248***	0.283***	0.270***	0.280***	0.226**	0.195*	2.302
		(0.030)	(0.031)	(0.031)	(0.030)	(0.086)	(0.095)	(2.426)
Grad Deg × Post Exp			-0.025	-0.026	-0.028^{+}	-0.032*	-0.023	-0.185
			(0.016)	(0.016)	(0.015)	(0.016)	(0.021)	(0.684)
Grad Deg × AFQT*				0.098**	0.022	0.117*	0.190**	4.643**
				(0.030)	(0.037)	(0.060)	(0.061)	(1.703)
Grad Deg × AFQT* × Post Exp					0.026**	0.026**	0.003	0.351
					(0.009)	(0.008)	(0.011)	(0.411)
Grad Deg × AFQT* × Pre Exp						-0.035	-0.072**	-1.589*
						(0.023)	(0.025)	(0.741)
Grad Deg × AFQT* × Pre × Post Exp							0.013^{+}	0.289
							(0.007)	(0.243)
R^2	0.430	0.453	0.454	0.455	0.455	0.454	0.454	0.404
Observations	114,167	114,167	114,167	114,167	114,167	114,167	114,167	114,167
Individuals	4,346	4,346	4,346	4,346	4,346	4,346	4,346	4,346
No. of GED/HS Degrees	.	100	100	100	100	100	100	100
No. of AA/BA Degrees	.	190	190	190	190	190	190	190
No. of Graduate Degrees	.	150	150	150	150	150	150	150
n_Gr	.	166	166	166	166	166	166	166

Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CHAPTER 3

OCCUPATIONAL SORTING, MULTIDIMENSIONAL SKILL MISMATCH, AND THE CHILD PENALTY AMONG WORKING MOTHERS

Disclamer

This chapter was co-authored with Gabrielle Pepin (pepin@upjohn.org). Gabrielle has approved that this work be included as a chapter in my dissertation.

3.1 Introduction

Despite considerable gender convergence in education and labor market experience during the late 1900s, substantial gender gaps in labor market outcomes remain in the United States and other industrialized nations (Blau and L. M. Kahn 2017; Goldin 2014; Goldin, Katz, and Kuziemko 2006). Parenthood plays a large role in explaining gender gaps, as researchers document large and persistent child penalties—gender gaps in labor market outcomes due to children—across countries and family structures.[1] Although the existence of the child penalty is well-established, there is limited evidence on its determinants (Angelov et al. 2016; Chung et al. 2017; Kleven, Landais, Posch, et al. 2019).

We contribute to the literature on understanding the sources of the child penalty by investigating the role of occupational sorting. As children limit the amount of time and energy parents may devote to their jobs, they likely affect labor market outcomes via occupation choice. A child penalty may arise if children lead women, who tend to bear the majority of child-rearing responsibilities even when they work full-time (Hsin and Felfe 2014), to sort into less-demanding, lower-paying occupations. Moreover, sorting into such "family-friendly" occupations may exacerbate the child penalty by worsening the match between women's skills and occupational requirements.[2]

In this paper, we first use the event-study approach proposed by Kleven, Landais, and Søgaard (2019) and data from the National Longitudinal Surveys of Youth (NLSY) to document child penalties among individuals born in the United States during the 1950s and 1960s and during the 1980s who later become working parents. By leveraging variation in the timing of first births, we allow fertility decisions to be

[1] See Angelov, Johansson, and Lindahl (2016), Chung et al. (2017), Cortes and Pan (2020), Kleven, Landais, Posch, et al. (2019), Kleven, Landais, and Søgaard (2019), Kleven, Landais, and Søgaard (2021), and Sieppi and Pehkonen (2019).

[2] See Bacolod and Blum (2010), Deming (2017), Guvenen et al. (2020), Heckman, Stixrud, and Urzua (2006), and Lise and Postel-Vinay (2020).

endogenous but assume that any unobservable determinants of labor market outcomes evolve smoothly around childbirth. Comparing effects across working mothers and fathers, we show that children cause long-run earnings gaps of over $200 per week in both the NLSY79 and NLSY97. While decreases in wages and hours worked among mothers generate child penalties in the NLSY79, increases in fathers' wages and hours generate penalties in the NLSY97.

We then use the event-study framework to estimate effects of children on occupational sorting. We show that in the NLSY79, children lead mothers to sort into lower-paying occupations in which employees tend to work fewer hours. While we do not observe this phenomenon among mothers in the NLSY97, occupational sorting still appears to play a large role in explaining the child penalty, as fathers sort into higher-paying occupations with higher average hours worked. Next, as occupational sorting likely affects the degree of complementarity between parents' skills and occupational requirements, we estimate effects of children on multidimensional occupation-skill mismatch. In other words, we estimate effects on the extent to which parents' math, verbal, science and mechanical, and social, or noncognitive, skills differ from those required by their occupations.

To examine mismatch, we proxy for skills using NLSY respondents' test scores on the Armed Services Vocational Aptitude Battery (ASVAB), along with information on sociability and extracurricular participation during childhood. We link these skill measures to O*NET, which documents occupational task content, and measure the relatedness of occupation tasks to skill categories using the United States Department of Defense's Defense Manpower Data Center crosswalk. Finally, we measure the distance between individuals' multidimensional skills and the importance of those skills in their occupations. We find that, among mothers in the NLSY79, children increase multidimensional occupation-skill mismatch by about 0.3 standard deviations, relative both to their own levels of mismatch from before birth and to those of fathers. In contrast, mothers in the NLSY97 exhibit decreases in skill mismatch post-childbirth. Nonetheless, large child penalties remain.

Our work falls at the intersection of existing literatures on the child penalty and occupational sorting. In terms of child penalties, there is extensive work on penalties in Europe, where earnings penalties range from about 20 percent in Denmark to about 60 percent in Germany.[3] Three papers study child penalties in the United States. Cortes and Pan (2020) and Kleven, Landais, Posch, et al. (2019) use data from the Panel

[3] See Angelov et al. (2016), Chung et al. (2017), Kleven, Landais, and Søgaard (2019), Kleven et al. (2021), and Sieppi and Pehkonen (2019).

Survey of Income Dynamics on parents who had first children during the late 1900s and early 2000s to document long-run penalties in earnings between 30 and 40 percent. Chung et al. (2017) find similar results using data from the Survey of Income and Program Participation linked to earnings records of parents whose first children were born between 1978 and 2011. They also document decreases in the size of the child penalty over time.

Regarding determinants of the child penalty, gender norms seem to matter, as Kleven, Landais, and Søgaard (2019) find that girls who grew up in families with traditional divisions of labor incur larger penalties when they become mothers. Additionally, researchers find that child penalties are larger among couples in which the father has more education, which is line with the theory of comparative advantage (Angelov et al. 2016; Chung et al. 2017). We contribute to the literature on understanding the sources of the child penalty by investigating the role of occupational sorting. Changes in workplace attributes due to children suggest that occupational sorting likely plays an important role in explaining child penalties. In particular, Kleven, Landais, and Søgaard (2019) find that, in Denmark, children lead women to move into workplaces that employ larger shares of women with children. They also find that children cause women to move from private into public sector employment, where work hours tend to be more flexible.

While we are the first to estimate effects of occupational sorting *due to children* on the gender gap, research on the extent to which occupational sorting explains gender gaps more generally dates back to the decompositional analyses of Blinder (1973) and Oaxaca (1973). More recently, Blau and L. M. Kahn (2017) estimate that occupational sorting explained about 10 percent of the gender wage gap in the United States in 1980 and about 30 percent of it in 2010. We improve upon traditional gender gap decompositions because we do not control for labor market choices likely affected by children, such as industry, in our analyses. Additionally, whereas children largely do not factor into traditional gender gap decompositions by construction, given similar numbers of children across men and women, we explicitly focus on the role of children in explaining gender inequality.

Finally, in studying effects of children on the degree of complementarity between parents' skills and occupations, we contribute to a growing literature on multidimensional occupation-skill mismatch (Addison, Chen, and Ozturk 2020; Guvenen et al. 2020; Lise and Postel-Vinay 2020; Speer 2017b). While much of the existing literature on multidimensional skills focuses exclusively on men, Addison et al. (2020), who study gender differences in skill mismatch over the lifecycle, is a notable exception. Using data from the NLSYs, the authors show that mothers exhibit greater mismatch in their occupations than men and childless women, though differences in mismatch have decreased over time. They also find that mismatch is relatively large among college-educated women and parents in occupations with more job flexibility. We build on Addison et al. (2020) by parsing the causal effects of children on mismatch from differences across mothers, fathers, and childless workers that may be unobservable to researchers.[4]

In the following section, we provide institutional details about access to parental leave and child care in the U.S. In Section 3.3, we describe the data. In Section 3.4, we estimate effects of children on labor supply and occupational sorting. In Section 3.5, we conclude.

3.2 Institutional Setting

Family policies in the U.S. are notoriously ungenerous. Before the enactment of the Family and Medical Leave Act of 1993 (FMLA), workers were not guaranteed any parental leave.[5] As of 1991, some 37 percent of female full-time workers in private-sector firms with at least 100 employees had access to unpaid leave, and only 2 percent had access to paid leave. Full-time male workers in similar firms were less likely to have access to leave: some 26 percent had access to unpaid leave, and 1 percent had access to paid leave (U.S. Bureau of Labor Statistics 2020).

Since 1993, firms with at least 50 employees must offer eligible employees 12 weeks of job-protected unpaid leave for childbirth or adoption.[6] To be eligible for leave, employees must have worked at the firm for at least 12 months and have accumulated at least 1,250 work hours. FMLA increased leave coverage substantially. In 1994, some 84 percent of full-time workers in private-sector firms with at least 100 employees had access to unpaid leave. Still, only 2 percent of full-time workers in similar firms had access to paid leave, and many workers were not covered by the law. Less than 50 percent of both part-time workers in firms with at least 100 employees and full-time workers in smaller firms had access to any leave benefits (U.S. Bureau of Labor Statistics 2020).

[4] Though the authors describe their work as taking a "descriptive approach rather than establishing definitive causal associations," in some specifications, Addison et al. (2020) use the age of the NLSY respondent's sibling at the time of the sibling's first birth as an instrument for the timing of the respondent's first birth to estimate effects of children on mismatch. We do not believe that sibling age at first birth passes the monotonicity assumption required of instruments and view Addison et al. (2020) as a valuable descriptive contribution to the literature.

[5] Some states mandated some type of parental leave before 1993, including five states that offer temporary disability insurance. The Pregnancy Discrimination Act of 1978 mandated that employers in states with temporary disability insurance treat pregnancy as a short-term disability, which allows mothers to receive partial earnings without job protection for around six weeks.

[6] Employers with at least 50 employees within 75 miles of the worksite for at least 20 weeks of the last year must offer 12 weeks of unpaid leave, though some states have lower firm size thresholds or require longer leave lengths. Leave also may be taken if the employee is in poor health or cares for a close relative who is in poor health. Employers may refuse job protection for their highest-paid 10 percent of employees if leave would generate economic harm.

While the U.S. still does not mandate paid parental leave, between 2004 and 2017, California, New Jersey, and Rhode Island implemented their own paid leave mandates. California and New Jersey offer six and Rhode Island offers four weeks of paid parental leave.[7] In each of these states, workers who meet a given work history requirement receive partial wage replacement up to a maximum weekly benefit, though only Rhode Island offers job protection beyond what is covered under FMLA. California and Rhode Island extend benefits to workers at small firms; New Jersey does not. As states enacted paid leave mandates, access to benefits increased: between 2005 and 2017, the proportion of private-sector workers with access to paid leave increased from 0.07 to 0.13. At the same time, the proportion with access to unpaid leave increased from 0.81 to 0.87 (U.S. Bureau of Labor Statistics 2020). Hence, while access to paid leave has increased in recent years, most workers still do not receive such benefits. Even mothers who have access to paid leave must return to work relatively shortly after childbirth to be guaranteed job protection.

In addition to limited parental leave benefits, the U.S. does not offer universal pre-kindergarten or child care, though universal schooling is available for children beginning at age five. Several states, however, operate their own universal pre-kindergarten programs for four- and, in some cases, three-year-old children. In particular, between 1995 and 2008, Florida, Georgia, Iowa, Oklahoma, Vermont, West Virginia, and Wisconsin implemented universal pre-kindergarten programs. Additionally, children aged three and four in families with incomes at or below the federal poverty level can participate in Head Start, a means-tested federal preschool program that was rolled out across the U.S. between 1965 and 1980. Head Start's objective is to promote children's cognitive and interpersonal development and school readiness through education, health and nutrition interventions, and family partnerships. An additional branch of Head Start, Early Head Start, was established in 1994 to serve pregnant women and children younger than age three who meet Head Start's income-eligibility criteria. Early Head Start consists of center-based care or home visits and is especially focused on nurturing healthy relationships between children and their caregivers.

Among households without access to free services, income support for early care and education is limited, though state and federal governments administer some cash benefits to families through the tax code. For instance, the Child and Dependent Care Credit (CDCC), which was introduced in 1976, subsidizes child care costs for working families. Between 2003 and 2020, households could claim up to \$3,000 in child care expenses per child for up to two children and receive CDCC benefits worth up to 35 percent of those expenses, or \$1,050. In addition, about 40 percent of workers can access dependent care flexible spending

[7] California increased the maximum leave length to eight weeks in 2020.

accounts (FSA) that their employers offer (U.S. Bureau of Labor Statistics 2020). Since 1986, employees who receive FSAs from their employers have been able to set aside up to $5,000 of earnings before taxes for dependent care expenses. The employer deducts this income from employees' paychecks, but employees are reimbursed for child care expenditures. Additional tax benefits for families with children, but not explicitly for child care, include the Child Tax Credit (CTC) and Earned Income Tax Credit (EITC). The CTC was introduced as a child benefit in 1997, and between 2003 and 2020, families could receive benefits of up to $1,000 per child. The EITC is an earnings subsidy targeted at low- and moderate-income families with children. As of 2020, maximum benefits for one-, two-, and three-child households were $3,584, $5,920, and $6,660, respectively.

Taken together, lack of access to family leave and universal early care and education services generates very high costs of child-rearing in the U.S. Under such constraints, the arrival of children may lead parents to move into less-demanding occupations that could negatively impact their long-run labor market outcomes. In light of this, we study effects of children on parents' labor supply and occupation choices in the following sections.

3.3 Data

To examine child penalties and occupational sorting, we link individual-level data on worker skills, occupations, and labor market outcomes from the NLSYs to occupational task content from O*NET.

3.3.1 NLSY79 and NLSY97

The NLSY79 and NLSY97 are nationally-representative panel surveys of individuals living in the United States aged 14 to 22 in 1979 and 12 to 16 in 1997, respectively. Biennial interviews of each cohort continue through the present, though interviews were conducted annually from 1979 to 1994 for the 1979 cohort and from 1997 to 2011 for the 1997 cohort. The NLSYs contain extensive information on individuals' demographics, family backgrounds, educational experiences, and labor market outcomes. Importantly for our study, the data document respondents' census occupation codes and months of birth for their children. Detailed information on individual characteristics and the long-panel nature of the surveys make the NLSYs well-suited to estimate long-run effects of children on labor market outcomes.

Another key advantage of the NLSYs is their inclusion of ASVAB test scores, which were administered to NLSY79 respondents in 1981 and to NLSY97 respondents in 1999. The ASVAB, which measures cognitive skills in ten subjects, was developed by the United States military in 1968, and since 1976, all military branches have used it to determine eligibility for military occupations. For example, to be eligible for a position as an electronics technician in the United States Navy, a recruit's composite score from the arithmetic reasoning, mathematics knowledge, electronics information, and general science sections of the ASVAB must exceed a certain threshold.[8] ASVAB scores likely are good proxies for individuals' cognitive skills, as military researchers show that scores on sections required for occupations predict job performance (Sims and Hiatt 2001; Welsh, Kucinkas, and Curran 1990).

To study child penalties and occupational sorting, we restrict the NLSY samples to working parents who fully transitioned into the labor market before having children.[9] Similarly to H. S. Farber and Gibbons (1996), Schönberg (2007b), and Speer (2017b), we designate an individual as having made a full transition into the labor market if they have not been enrolled in school for two consecutive years and have worked for at least 30 hours per week in at least half of the weeks during those two years. We exclude from the sample individuals with fewer than two years of labor market experience before childbirth, individuals with missing ASVAB scores, active-duty military and veterans, and the self-employed. We classify an individual as a working parent if they worked for at least ten hours per week at a main employer for some period during the year of their first childbirth, the year immediately preceding childbirth, the year immediately following childbirth, and at least half of the following four years.[10] As in Cortes and Pan (2020), we require that individuals complete at least one interview before childbirth, at least one interview after childbirth, and at least four interviews during the sample period.

3.3.2 O*NET and the Defense Manpower Data Center Crosswalk

O*NET, a database maintained by the United States Department of Labor, documents knowledge, skills, and abilities—henceforth, "tasks"—required of occupations. Specifically, for each occupation in the Standard Occupational Classification (SOC) system, expert job analysts, job supervisors, or job incumbents assign scores for the importance of 277 tasks.[11] As in Addison et al. (2020) and Guvenen et al. (2020), we use

[8] See `https://www.military.com/join-armed-forces/asvab` for information on ASVAB score requirements for military occupations.

[9] We include evidence on child penalties among all parents in the online appendix. In constructing the sample, we include the NLSY79 and NLSY97 Black and Hispanic oversamples but exclude the NLSY79 economically disadvantaged oversample, which was discontinued in 1991. Results from analyses in which we exclude the Black and Hispanic oversamples are similar and available upon request.

[10] Results are robust to alternative definitions of working parents and are available upon request.

[11] We map the SOC occupations into occupation categories included in the NLSYs using census occupation codes.

the Defense Manpower Data Center (DMDC) crosswalk, which was created by the United States Department of Defense, to measure the relatedness of O*NET tasks to ASVAB section scores. The crosswalk includes information on twenty-six O*NET tasks to which personnel research psychologists assign "at least a moderately strong probability" of being related to at least one of the following ASVAB section tests: word knowledge, paragraph comprehension, arithmetic reasoning, mathematics knowledge, general science, mechanical comprehension, and electronics information (ASVAB Career Exploration Program 2011). For each of the tasks, psychologists and psychometricians assign relatedness scores to each of the aforementioned ASVAB section tests.

3.3.3 Creating Multidimensional Occupation-Skill Mismatch Measures

Figure 1 illustrates the procedure for creating the multidimensional occupation-skill mismatch measure for a given individual and occupation. First, we create measures of the relevance of ASVAB section scores to occupational requirements. This allows us to link occupational requirements to individuals' bundles of skills, as demonstrated on the ASVAB test. To create measures of ASVAB score relevance to a given occupation, we use the measures of task importance within the occupation from O*NET and the measures of task relatedness to ASVAB section test scores from the DMDC crosswalk.[12] Focusing on the twenty-six tasks included in the crosswalk, we multiply the vector of occupational task importance measures by the matrix of ASVAB task-relatedness measures. This yields measures of ASVAB section score relevance to the given occupation, which we normalize to have standard deviations of 1.

Next, as each military branch combines ASVAB section scores in different ways to determine recruits' suitability for occupations, we follow Addison et al. (2020) in creating four skill categories—math, verbal, science/mechanical, and social.[13] The first three categories correspond to sections of the ASVAB test. Specifically, the mathematics knowledge and arithmetic reasoning sections of the ASVAB correspond to the math category; the word knowledge and paragraph comprehension sections correspond to verbal category; and the general science, mechanical comprehension, and electronics information sections correspond to science/mechanical category. Because multiple ASVAB sections comprise each of the skill categories, as shown in Figure 3.1, we follow Guvenen et al. (2020) and apply principal component analysis (PCA)

[12]Because the scale of the DMDC relatedness score is somewhat arbitrary, we rescale each ASVAB section's twenty-six task-relatedness scores to sum to 1.

[13]Results are robust to using alternative skill categories proposed by Addison et al. (2020), Guvenen et al. (2020), Lise and Postel-Vinay (2020), and Speer (2017b) and are available upon request.

to both the vector of occupational relevance measures and to the vector of the individual's ASVAB section scores.[14] In doing so, for each skill category, we create a measure equal to the first principle component of the pertinent relevance measures or ASVAB section scores. For example, the verbal relevance measure is the first principle component of the relevance of the word knowledge and paragraph comprehension ASVAB sections to the given occupation. Analogously, the verbal skill measure is the first principle component of the individual's scores on the word knowledge and paragraph comprehension sections of the ASVAB. We then scale the occupational skill relevance measures and the individual's skill measures into percentile ranks among occupations and NLSY respondents in their cohort, respectively.[15]

In addition, we create a social skill category. To construct a measure of the relevance of social skills to a given occupation, we follow Addison et al. (2020), Guvenen et al. (2020), and Deming (2017) and rely on the following occupational task-relatedness measures from O*NET: social perceptiveness, coordination, persuasion, negotiation, instructing, and service orientation. To construct a social skill measure for each individual, we again follow Deming (2017) and use information on self-reported sociability during childhood and adolescence and club and sport participation during high school available in the NLSYs. For both the task-relatedness and skill measures, we scale the standard deviation of each of the components to equal one and apply PCA. As with the cognitive skill measures, we then convert the social occupational task-relatedness and skill measures into percentile ranks.

Finally, we compare individuals' bundles of skills to those required by their occupations. We follow Guvenen et al. (2020) and define multidimensional occupation-skill mismatch as follows:

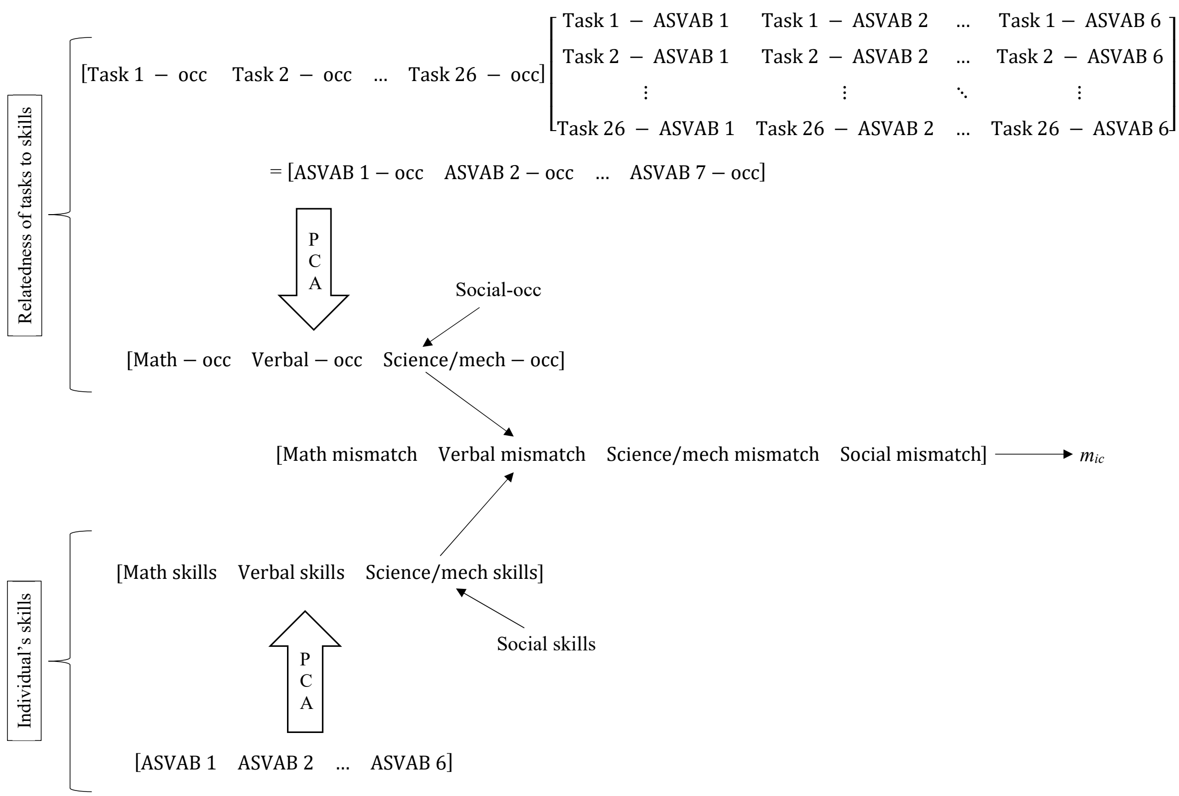
$$m_{ic} = \sum_{l=1}^{4} w_l * |q(a_{il}) - q(r_{cl})|, \tag{3.1}$$

where a_{il} is the skill measure for individual i in skill category l. r_{cl} represents the relevance of skill l within occupation c. $q(a_{il})$ and $q(r_{cl})$ denote the corresponding percentile ranks of individual skill and

[14] We follow Altonji, Bharadwaj, and Lange (2012b) in standardizing ASVAB section scores to account for differences in age at the time the test was administered and test format, as the NLSY79 cohort took a pencil and paper version of the ASVAB, and the NLSY97 cohort took a computer-assisted version of the test. To adjust for test format differences, we use a crosswalk based on scores of individuals randomly assigned to one of the two test formats (Segall 1997). We then perform an equipercentile mapping to age 16 separately for each NLSY cohort. In other words, we assign test scores of those who took the test at age a and scored in the qth percentile among age a test takers the corresponding qth-percentile score of those who took the test at age 16. In doing so, we assume that the relative ranking of an individual's score in their cohort's score distribution does not depend on when the cohort took the test. We also assume that the level of skill associated with a score in the qth percentile of the age a score distribution is the same as that associated with a qth-percentile score in the age 16 score distribution. We do not restrict scores across NLSY cohorts and normalize them to have a mean of 0 and standard deviation of 1 in 1979.

[15] When we scale occupational skill relevance measures, we weight each occupation by the number of individuals engaged in it in the individual's NLSY cohort.

Figure 3.1: Procedure for Creating Multidimensional Occupation-Skill Mismatch Measure



Notes: Procedure for creating multidimensional occupation-skill mismatch measure for a given individual and occupation. "Task – occ" is the O*NET score of the importance of a task for a given occupation. "Task – ASVAB" is the DMDC crosswalk's relatedness score of a task for a given ASVAB section test.

occupational relevance. w_l is the first principle component of $\sum_{l=1}^{4} |q(a_{il}) - q(r_{cl})|$. For ease of interpretation, we standardize m_{ic} to have a standard deviation of 1.

Furthermore, we use data on individuals' skills and those required by their occupations to study skill overmatch and undermatch, or the extent to which individuals are overqualified or underqualified for their occupations, respectively. We define skill overmatch as follows:

$$om_{ic} = \sum_{l=1}^{4} \mathbb{1}[(q(a_{il}) - q(r_{cl})) > 0](w_l * (q(a_{il}) - q(r_{cl}))), \tag{3.2}$$

where the variables are the same as those listed in Equation (3.1). Equation (3.2) implies that overmatch increases with the difference between an individual's skills and those required by their occupation. If none of an individual's skill measures exceed those required by their occupation, then the individual is not overqualified for the occupation, and om_{ic} equals 0. Analogously, we define skill undermatch:

$$um_{ic} = \sum_{l=1}^{4} \mathbb{1}[(q(a_{il}) - q(r_{cl})) < 0](w_l * (q(a_{il}) - q(r_{cl}))). \tag{3.3}$$

We standardize both om_{ic} and um_{ic} to have standard deviations of 1.

3.3.4 Current Population Survey Outgoing Rotation Groups

Finally, we use data from the Current Population Survey (CPS) to estimate average weekly earnings and hours worked by occupation and decade. The CPS, a nationally-representative monthly survey of over 65,000 households, is designed to measure employment. Households in the survey are interviewed for four months, ignored for eight months, then interviewed for four more months. Since 1982, respondents have documented labor force status, occupation, and usual hours worked per week during each month in the survey. We use data from the fourth and eighth outgoing interviews, during which respondents also report usual weekly earnings. Specifically, to create measures of average earnings and hours worked within occupations held by NLSY respondents, we estimate regressions of the given labor market outcome on census occupation dummies using sample weights. To allow for changes in average earnings by occupation over time, we estimate separate regressions for the 1980s, 1990s, 2000s, and 2010s.

3.3.5 Summary Statistics

Table 3.1 displays summary statistics for working mothers and fathers separately by NLSY cohort as of the year before their first childbirth. At that time, working parents in the NLSY79 tend to be about 25 years old; working parents in the NLSY97 are closer to 24 years old on average. About 50 percent of parents in the NLSY79 are married during the year preceding childbirth, whereas only about 40 percent of parents in the NLSY97 are married. Across both cohorts, pre-childbirth average weekly earnings (2010 dollars) tend to be higher for fathers than for mothers, but the gender gap shrinks across cohorts, as gender differences in average weekly earnings decrease from \$139 in the NLSY79 to \$60 in the NLSY97. Both increases in mothers' average earnings and decreases in fathers' average earnings drive the shrinking earnings gap. Similarly, fathers' average hourly wages decrease from \$15.72 to \$15.51 across cohorts while mothers' average hourly wages increase from \$13.84 to \$14.47. Fathers tend to work more hours per week than mothers, and average weekly hours worked decrease across cohorts for both genders: fathers' average hours decrease from 43 to 40, and mothers' from 39 to 38.

Table 3.1: Summary Statistics

	NLSY79		NLSY97	
	Fathers	Mothers	Fathers	Mothers
Age	25.15	25.06	24.32	24.24
	(3.65)	(3.79)	(3.20)	(3.38)
White	0.485	0.477	0.717	0.721
	(0.500)	(0.500)	(0.451)	(0.449)
Black	0.267	0.275	0.188	0.152
	(0.443)	(0.447)	(0.391)	(0.360)
Hispanic	0.248	0.248	0.095	0.127
	(0.432)	(0.432)	(0.294)	(0.333)
Married	0.464	0.535	0.376	0.418
	(0.499)	(0.499)	(0.485)	(0.494)
Weekly earnings ($)	683	544	632	572
	(493)	(414)	(403)	(452)
Hourly wage ($)	15.72	13.84	15.51	14.47
	(9.65)	(10.54)	(9.48)	(9.69)
Weekly hours	43.15	38.69	40.22	37.76
	(10.61)	(8.96)	(9.64)	(10.23)
Avg weekly earnings within occupation ($)	721	599	594	492
	(274)	(243)	(216)	(206)
Avg weekly hours within occupation	38.55	35.52	35.84	32.55
	(4.14)	(4.59)	(4.42)	(4.14)
Occupation-skill mismatch	2.03	1.75	2.06	2.00
	(1.076)	(0.927)	(0.986)	(0.997)
Overmatch	0.719	0.557	0.413	0.310
	(0.657)	(0.553)	(0.435)	(0.395)
Undermatch	–0.389	–0.429	–0.670	–0.721
	(0.507)	(0.497)	(0.587)	(0.584)
Observations	875	622	399	433

Note: Summary statistics among working parents in the NLSY79 and NLSY97 as of the year before their first childbirth. Standard deviations are listed in parentheses.
Source: Authors' calculations using the NLSY79, NLSY97, and Current Population Survey.

Turning to occupation choice, Table 3.1 shows that fathers tend to sort into occupations in which, on average, employees work more hours and garner more earnings, compared to mothers. In the NLSY79, average earnings within occupation are about $700 and $600 for working fathers and mothers, respectively. Average earnings within occupation are about $600 for fathers and $500 for mothers in the NLSY97. Despite sorting into lower-paying occupations, mothers tend to exhibit better occupation-skill matches. Average mismatch measures, denoted by m_{ic} in Equation (3.1), are 1.75 and 2.00 for mothers in the NLSY79 and NLSY97, respectively. This compares to 2.03 for fathers in the NLSY79 and 2.06 for fathers in the NLSY97. Mothers also exhibit less skill overmatch and more skill undermatch compared to fathers. Measures of overmatch and undermatch indicate that, based on their skills and those required by their occupations, both mothers and fathers exhibit lower degrees of overqualification and higher degrees of underqualification for their occupations over time.

3.4 Evidence on Occupational Sorting and the Child Penalty

3.4.1 Empirical Strategy

We use the event-study method proposed by Kleven, Landais, and Søgaard (2019) to estimate effects of children on parents' labor supply and occupation choices. In doing so, we allow for endogenous fertility but assume that unobservable determinants of labor market outcomes evolve smoothly around childbirth. Under this assumption, we attribute any discontinuity in outcomes around childbirth to effects of children. The smoothness assumption would be violated if, for instance, parents time childbirth to coincide with a job promotion. We estimate the following event-study model separately by gender and NLSY cohort, where event time $t = 0$ during the year the individual has their first child:

$$Y^g_{ist} = \sum_{j \neq -1} \alpha^g_j * \mathbb{1}[j = t] + \sum_k \beta^g_k * \mathbb{1}[k = age_{is}] + \sum_y \gamma^g_y * \mathbb{1}[y = s] + \delta^g X^g_i + v^g_{ist}. \tag{3.4}$$

Y^g_{ist} is the outcome of interest for individual i of gender g in year s relative to event time t. We omit the indicator for $t-1$ so that the $\hat{\alpha}^g_j$ coefficients measure effects of children relative to the year before birth. We include age and year dummies to control non-parametrically for life-cycle and time trends, such as inflation and business cycles. X^g_i includes controls for education at labor market entry and race and an indicator for entering the sample after $t-5$.[16] We cluster standard errors at the individual level.

Equation (3.4) leverages differences in birth timing, conditional on age, year, and individual characteristics, to estimate post-childbirth effects of children on labor market outcomes. Short-run measures of the child penalty capture effects of a first child, whereas long-run measures may capture effects of total fertility. Despite differences in interpretation of effects, for both short- and long-run penalties, $\hat{\alpha}^g_j$ only captures effects of children that are realized after childbirth. Thus, to the extent that children reduce pre-childbirth labor market investments of women relative to men, we underestimate effects of children on the gender gap.

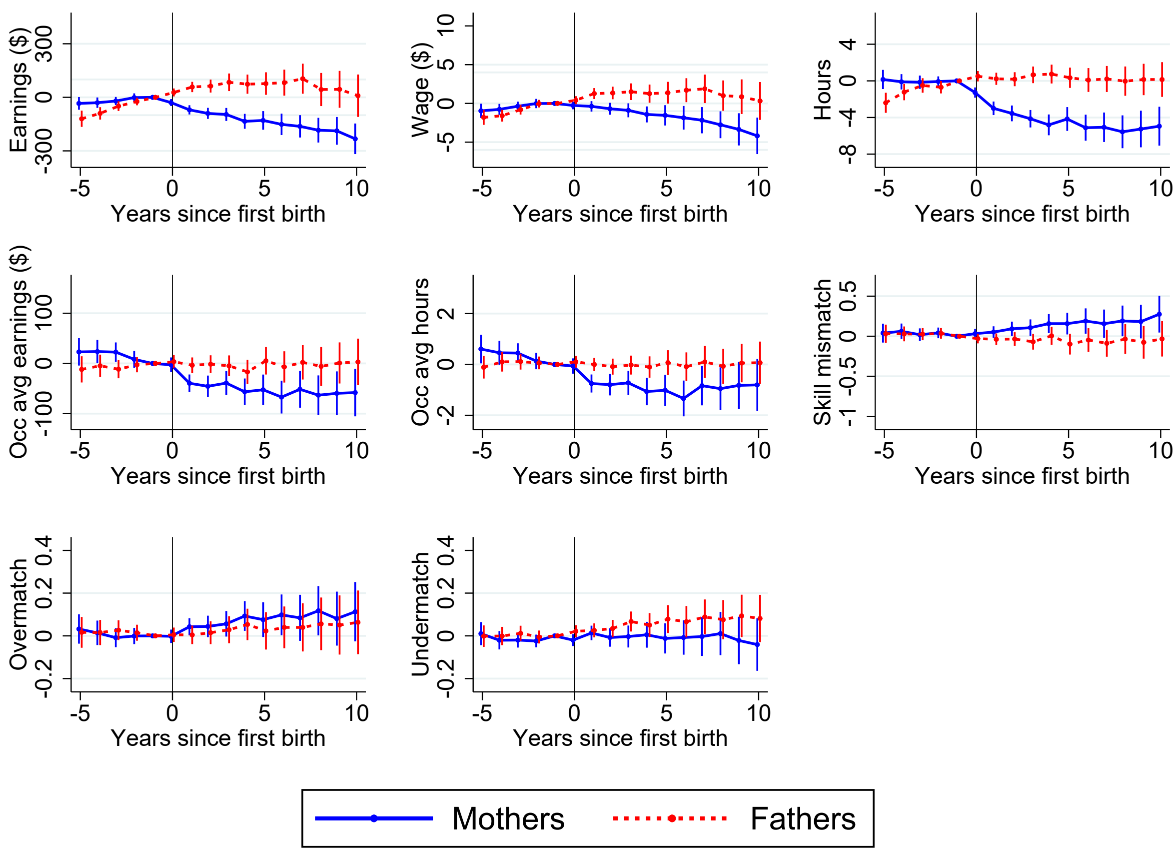
3.4.2 Results

Figure 3.2 displays results among working mothers (solid blue lines) and fathers (red dashed lines) in the NLSY79. The upper-left panel of Figure 2 documents effects on weekly earnings. The panel shows that,

[16] We include controls for individual characteristics in our models, whereas Kleven, Landais, and Søgaard (2019) generally do not, because our data exhibit incomplete overlap in event time, age, year, and education and in event time, age, year, and race. Thus, if we did not control for individual characteristics, the age and year dummies would capture effects of some but not all education levels and races.

conditional on life-cycle and time trends and time-invariant individual characteristics, pre-birth earnings trajectories are fairly similar for mothers and fathers. Then, when children arrive, mothers' and fathers' earnings paths diverge. Mothers experience immediate decreases in earnings that continue to grow for at least ten years after childbirth. Fathers' earnings increase initially but settle around prebirth levels by ten years post-childbirth, though 95 percent confidence bands indicate that long-run estimates are a bit noisy. Results imply that children cause long-run earnings gaps of over $200 per week between working mothers and fathers.

Figure 3.2: Effects of Children in the NLSY79



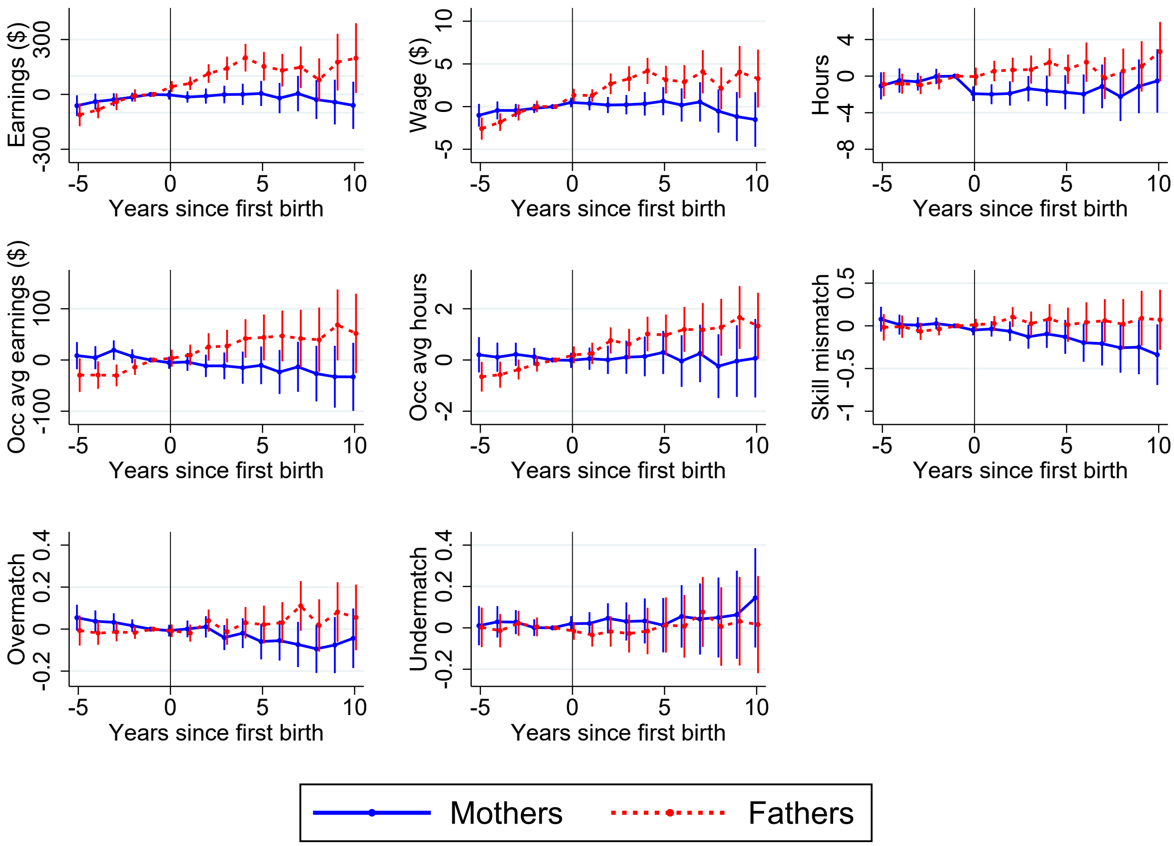
Notes: Effects of children on weekly earnings; hourly wages; weekly hours worked; average weekly earnings within occupation; average weekly hours worked within occupation; and standard deviations of multidimensional occupation-skill mismatch, skill overmatch, and skill undermatch among working mothers (blue solid line) and fathers (red dashed line) in the NLSY79. Vertical lines denote 95% confidence bands.

Similarly, the upper-middle and right panels of Figure 3.2 show effects on hourly wages and hours worked per week, respectively. As with earnings, working parents' wages and hours worked trend fairly similarly before birth but diverge immediately thereafter. Mothers' hours worked continue to decrease until they plateau at around 4 fewer hours per week five years post-birth. Mothers' wages continue to fall through at least ten years post-birth, when wages are nearly $5 less per hour. This constitutes a 29 percent decrease from the pre-birth mean. Meanwhile, fathers' hours worked remain relatively constant, and their wages increase slightly, at least through the medium-run.

The remaining panels of Figure 3.2 display effects on occupational sorting. Results imply that while fathers remain unaffected, the arrival of children immediately causes women to sort into occupations with lower average earnings in which employees tend to work fewer hours per week. Effect sizes, which remain relatively stable over time, imply that children lead women to enter occupations that pay about $50 less per week and where employees work about 1 fewer hour per week on average. Turning to effects on skill mismatch, children, again, do not seem to affect fathers' outcomes. Mothers' mismatch gradually increases after the arrival of children, however. Ten years after birth, their mismatch has increased by about 0.3 standard deviations. Results from the lower panels of Figure 2 suggest that increases in mismatch are driven by mothers becoming more overqualified for their occupations.

Next, Figure 3.3 presents results among working parents in the NLSY97. As with results among parents in the NLSY79, the arrival of children generates gender gaps in earnings, wages, and hours worked. Unlike in the NLSY79, however, better labor market outcomes among fathers, rather than worse outcomes among mothers, tend to drive results. Specifically, fathers' earnings, wages, and hours worked continuously increase through four years post-childbirth. At the same time, mothers experience very little change in earnings and wages and work about two fewer hours per week. Thus, while the sizes of gender gaps in labor market outcomes remain similar across cohorts, fathers' responses to children drive child penalties in the more recent cohort. In line with this, children lead fathers to sort into higher-paying occupations that require more hours worked per week while average earnings and hours worked within mothers' occupations do not change. Nonetheless, children lead mothers, but not fathers, to sort into occupations that are better matches for their bundles of skills. As estimated effects on overmatch and undermatch are rather noisy, it is unclear whether children lead mothers to move into occupations for which they are overqualified or underqualified.

Figure 3.3: Effects of Children in the NLSY97



Notes: Effects of children on weekly earnings; hourly wages; weekly hours worked; average weekly earnings within occupation; average weekly hours worked within occupation; and standard deviations of multidimensional occupation-skill mismatch, skill overmatch, and skill undermatch among working mothers (blue solid line) and fathers (red dashed line) in the NLSY97. Vertical lines denote 95% confidence bands.

3.5 Conclusion

Consistent with existing literature, we document sizable long-term child penalties among working parents in the NLSY79 and NLSY97.[17] Evidence suggests that occupational sorting plays an important role in explaining child penalties, as children lead mothers, but not fathers, in the NLSY79 to sort into occupations with lower average pay in which employees tend to work fewer hours per week. In the NLSY97, children cause fathers, but not mothers, to sort into higher-paying occupations with higher average hours worked. Additionally, children generate both absolute and relative increases in multidimensional occupation-skill mismatch among mothers in the NLSY79. Evidence suggests that mismatch effects are driven by mothers

[17] See Angelov et al. (2016), Chung et al. (2017), Cortes and Pan (2020), Kleven, Landais, Posch, et al. (2019), Kleven, Landais, and Søgaard (2019), Kleven et al. (2021), and Sieppi and Pehkonen (2019).

becoming more overqualified for their occupations. In contrast, mothers in the NLSY97 exhibit decreases in skill mismatch post-childbirth, but better occupation matches do not compensate for increases in wages and hours worked among fathers, and large child penalties remain.

While we followed existing literature in creating measures of multidimensional occupation-skill mismatch, future research may investigate whether other measures of mismatch better predict wages, conditional on an individual's skills and occupational requirements. For example, there may be nonlinear returns to skill mismatch across the distribution of mismatch levels, or there may be asymmetric returns to overmatch and undermatch. In addition, creating a single multidimensional occupation-skill mismatch measure could mask a considerable amount of heterogeneity in mismatch across skill categories. Thus, in future work, researchers may propose various measures of skill mismatch and empirically test their efficacy.

Perhaps most importantly, future research may further investigate drivers of occupational sorting and their effects on gender inequality, as well as differences in outcomes over time. For instance, researchers may examine heterogeneity in outcomes across demographic groups and individuals with different preferences over work and family during adolescence. Additionally, descriptive evidence on occupation flows and associated changes in skill requirements could provide further evidence on the channels through which occupational sorting occurs. Finally, examining heterogeneity in outcomes across groups with and without access to family-friendly work policies, such as paid leave and universal pre-K, may inform policymakers about the effectiveness of such policies in reducing the child penalty and gender wage gap.

APPENDICES

APPENDIX A

CHAPTER 1 APPENDIX

A.1 Data Appendix

The data used to study the empirical predictions discussed in the previous section come from the 2017 release of the National Longitudinal Survey of Youth 1997 cohort (NLSY97), a nationally representative survey of 8,984 men and women who were between the ages of 12 and 16 on December 31st, 1996. This survey, which was designed to capture the evolution of employment career paths of individuals from the time they leave school through adulthood, was administered annually between 1997 and 2011 before switching to its current biennial format following the completion of the 2011 interview round. During these interviews, extensive events histories are collected from the respondents related to a variety of topics covering employment, program participation, and education, as well as other important life events such as marital status and parental cohabitation. The event history data available in the NLSY97, as well as the detailed employer roster, make this an ideal source from which to study the effects of early career job loss.[1]

There are two additional features of the NLSY97 data that make it ideal for this type of study. The first of these features is the detailed respondent report regarding the reason for job loss. That is, the data allow me to identify whether a worker lost his or her job due to layoff, plant closure, or termination with cause. This stands in contrast to other data sources that do not allow for termination with cause to be separately identified from layoffs (PSID), or that do not offer detailed information about such terminations at all (DWS). The second key feature of the NLSY97 is the information available on pre-market skills, aptitude and cognition tests (such as the ASVAB), as well detailed histories on a variety of topics such as incarceration or drug use, which can be used as proxies for ability/quality in empirical work. For this study, I will be using the age adjusted AFQT scores created by Altonji et al. (2012a) that are directly comparable to the AFQT measures used in FG and AP. AFQT scores, which are derived from the math and reading portions of the ASVAB test, have been a standard measure of a worker's productive ability used in the literature dating back to Neal and Johnson (1996). For ease of interpretation, I normalize the AFQT scores to be mean zero with unit variance in the sample.

[1]The employer roster provides information on a variety of employment characteristics for each job an individual works ranging from the industry/occupation to whether the worker enjoys his/her job. Additionally, each employer on the roster is assigned a unique identifier which can be matched with the employer identifiers used in the event history data.

To study these effects, I match the employer identification numbers in the NLSY97's weekly employment array with information from the employer roster to construct an aggregated quarterly employment panel for each individual which details the total earnings and hours worked across up to five reported employers in quarter. From this, I identify each worker's primary employer in a quarter as the employer for which the worker worked the greatest number of hours, deferring to the employer with the longest tenure in the event of a tie. In the analysis that follows, I will be focusing on the wage effects of job loss based on the hourly wages for each of these primary employers, which is free from potential confounding effects that could be present if the average wage across multiple employers is used. One of the main benefits of a quarterly panel, as opposed to a yearly panel commonly used in studies of young worker displacement (e.g., Stevens 1997; Kletzer and Fairlie 2003), is the level of precision it affords in examining the wage dynamics of these displaced workers over time, especially given the focus on young workers who tend to experience more dramatic earnings/wage growth over relatively short periods of time due to job mobility, career advancement, etc., relative to older workers. Further, this level of aggregation allows me to construct a measure of labor market experience that evolves at the quarterly level and is more precise than conventional measures of potential labor market experience that must be used in the absence of work history data, as well as those that rely solely on a respondent's self-reported labor market experience which is subject to an evolving recall bias that grows as the respondent ages.

Involuntary job loss measures are defined based off of each respondent's self reported reason for why they are no longer employed with a previously identified primary employer. To avoid issues related to temporary or seasonal work, I ignore jobs that end before the respondent can accumulate at least 13 weeks (three months) of tenure. Any individuals who report a reason for job separation to be due to layoff, plant closure, or discharged/fired are considered to have involuntarily lost their job and thus make up my job loss sample. This sample is further broken down into three categories based on the type of job loss the respondent experienced, specifically layoff, plant closed, or fired. As the primary focus of this paper is on early career job loss, the analysis that follows will focus on the first involuntary job loss experienced by an individual as these are more likely to occur early in a worker's career and it avoids issues that arise when trying to separate the effect of a second job loss from the first as these are often correlated events.[2]

Finally, in order to analyze the relationship between employer learning and job loss through wage

[2] Stevens (1997) and Michaud (2018) note that the probability of involuntarily losing a job increases substantially if a worker has already experienced a prior job loss.

changes, I restrict my sample to individuals who have made their first long-term transition from school to the labor market prior to 2008. This transition is defined as being the first quarter in which an individual does not increase their education level the following year and will have worked at least 30 hours per week for half of the weeks during the following two years. This labor market entry definition is comparable to the one used in FG, though alternative definitions will also be considered for robustness checks. Additionally, for each worker, actual experience is defined as the number of weeks an individual has worked at least 30 hours divided by 50, and potential experience is defined as the number of quarters since an individual began their first employment spell divided by four. See Appendix A.1 for further information regarding the definition of a worker's first long-term transition into the labor market, as well as a complete description of the sample construction process and the methods used to create each of the relevant variables used in the study.

A.1.1 Identification of Layoff/Job Loss Sample

I identify workers who have experienced some form of involuntary job loss using the individual's self reported reason for why they are no longer employed with a previously reported employer. To avoid issues related to temporary or season work, I ignore jobs that end before the respondent can accumulate at least 13 weeks (three months) of tenure. Any individuals who report a reason for job separation to be due to layoff, plant closure, or discharged/fired are considered to have involuntarily lost their job and thus make up my job loss sample. Within this job loss sample, I further identify each type of job loss specifically, as well as the portion of the sample that would be considered "displaced" (i.e. lost their job due to layoff or plant closure) in the traditional job displacement literature. As the primary focus of this paper is on early career job loss, the analysis that follows will focus on the first involuntary job loss experienced by an individual as these are more likely to occur early in a worker's career and it avoids issues that arise when trying to separate the effect of a second job loss from the first as these are often correlated events.[3]

A.1.2 Additional Summary Statistics

Additionally, workers in the layoff sample are far more likely to have started their labor market career in a construction, production, or transportation related occupation than those in the non-job loss sample, while being far less likely to have started their labor market careers in a management or other professional

[3] Stevens (1997) and Michaud (2018) note that the probability of involuntarily losing a job increases substantially if a worker has already experienced a prior job loss.

occupation. Looking specifically at the laid-off worker sample, these individuals are far less likely to be female (35 percent compared to 46-54 percent for the other samples), far more likely to have started their labor market career in a construction related occupation,

Table **??** reports additional descriptive statistics for several of the key variables over the entire sample period. As would be expected, workers in each of the job loss samples have lower average career wages than the average for workers in the the non-job loss sample. Perhaps more surprisingly, despite having at least a half year or more potential experience on average, workers in each of the job loss samples average the same, or less, actual experience relative to workers in the non-job loss sample, indicating that some workers may have had some difficulty returning to work following their job loss. Also of interest is that workers in each of the job loss samples gained around the same amount of additional education after initially entering the labor market as workers in the control sample at around half a year of additional education on average. While we may be concerned the workers in the job loss sample were more likely to re-enroll in school following their job loss, at least descriptively, the workers in the job loss sample do not appear to be disproportionately increasing their education as a means of recovering from their job loss, though I will explore this in more depth in the main empirical analysis.

A.2 Creation of Residual AFQT Scores

To create an ability measure that is orthogonal to the information available to prospective employers when a worker first enters the labor market, I create a sample consisting of each individual's first period in the labor market (i.e. the first period they are employed and report a non-zero wage). Then, following FG, I define residual ability z_i^* as,

$$z_i^* = z_i - E^*(z_i \mid X_{i0}, \omega_{i0}),$$

where E^* is the linear projection of an ability measure z on a vector of observable characteristics X_{i0} and the worker's first period wage ω_{i0}. FG show that z^* is equivalent to observing employers' expectation error in a worker's ability for experience levels $t > 0$, and can be used by researchers to assess the effects of employer learning.[4] In practice, I regress each worker's AFQT score on a vector of observable characteristics and their first period wage, and then use the fitted values from this regression to calculate each worker's residual AFQT score.

The vector of observable characteristics used to create the residual AFQT scores contains five education dummy variables ($<$ 12 years, 12 years, 13-15 years, 16 years, and 17+ years), an indicator for part-time status, the interaction of part-time status and each education dummy, indicators for race, sex, marital status, marital status interacted with sex, age in years ($<$18, 18-19, 20-21, 22-23, 23-24, 25+), birth year, current year and quarter, and the log of the worker's wage. Additionally, I include indicators for the number of employees at each worker's employer ($\leq$10, 11-50, 51-100, and 100+), as well as the interaction of each of these indicators with each worker's log wage. This is done to account for potential differences in an employer's ability to judge a worker's true ability based on the employer's size, and to address a number of recent studies which have found that the size of a worker's first employer is an important determinant of labor market career trajectories (e.g., Moscarini and Postel-Vinay 2012; Arellano-Bover 2020).

This regression accounts for roughly 40 percent of the variation in AFQT scores, which is noticeably lower than the R^2 value found in FG (53 percent of variation accounted for using their NLSY79 sample

[4] Light and McGee (2015b) use a slightly different approach than FG. They regress their z measures only on the observable characteristics used in their model, leaving out the entry period wage. The advantage of that approach is that it does not require the entry period wage to be dropped in log-wage regression models, while still purging their ability measure of any correlation to observable characteristics. However, that approach does not purge the correlation between characteristics that are only observed by the employer and the ability measure, which will complicate the interpretation of the estimated return to ability over experience, since certain aspects of learning will be correlated with these observable characteristics, meaning their $\hat{z}$ is a biased measure of employers' true expectation errors. This issue is especially problematic in this context as the characteristics observable to employers are likely to have some form of correlation with the any job loss signal, which would make it impossible to distinguish between the signaling effect of the job loss and this correlation.

and nearly identical controls). This seems to be in line with recent empirical evidence from Altonji et al. (2012a), who find that the ability distribution has widened over time, and that demographic characteristics, such as race and gender, appear to play a less predictive role in an individual's AFQT score among the more recent cohort.[5] Lastly, while $AFQT_i^*$ is mean zero by construction, to make the measure comparable to the standardized AFQT scores, it is normalized to have unit-variance.

[5] It is also possible that this smaller relationship between observable characteristics and AFQT scores could be related to the recent evidence that the return to cognitive ability has generally been decreasing over the past few decades (e.g., Castex and Dechter 2014; Beaudry et al. 2016), though evidence from other studies using more updated data, such as Ashworth et al. (2020), among others, suggest that this may not be the case.

APPENDIX B

CHAPTER 2 APPENDIX

B.1 Traditional Model of Employer Learning

The traditional employer learning model is derived in the following way. Following AP's notation, let y_{it} denote the log productivity of worker i with t years of experience as:

$$y_{it} = rs_i + \alpha_1 q_i + \eta_i^* + \tilde{H}(t), \tag{B.1}$$

where η_i^* denotes the total transferable productive ability of worker i that is not directly observed by either the labor market or the econometrician; s_i denotes a vector of time-invariant characteristics of worker i that are observed by both the labor market (current and prospective employers) and the econometrician; q_i denotes a vector of time-invariant characteristics of worker i that are observed by the market, but not the econometrician; and $\tilde{H}(t)$ is the experience profile of productivity, with $\tilde{H}(0) = 0$.

In addition to the above, let z_i denote a vector of time-invariant correlates of worker i's productive ability that are not observed by employers, but are observed by the econometrician (e.g. AFQT score). Following FG, I assume that z_i has no direct effect on output, conditional on s_i, q_i, and η_i^*. Then, define η_i such that $\eta_i = \eta_i^* - \beta z_i$, where $\beta z_i = E(\eta_i^* | z_i)$, as the portion of a worker's total transferable productive ability that is orthogonal to z_i. For convenience, and to match the analysis of AP, assume that z_i is scaled such that β is always positive. Additionally, in much of the analysis below I suppress the i subscript.

Assume that the market's conditional expectations of η and z are linear in both s and q. Then, when $t = 0$, we can decompose η and z as follows:

$$\begin{aligned} z &= E(z|s,q) + \tilde{z} = \gamma_1 q + \gamma_2 s + \tilde{z} \\ \eta &= E(\eta|s,q) + \tilde{\eta} = \alpha_2 s + \tilde{\eta}, \end{aligned} \tag{B.2}$$

where $\tilde{\eta}$ and $\tilde{z}$ represent the portions of these ability measures that are uncorrelated with characteristics observed by the market when a worker first enters the labor market. Substituting this decomposition into Equation (B.1) yields:

$$\begin{aligned} y_t &= (r + \alpha_2 + \beta\gamma_2)s + (\alpha_1 + \beta\gamma_1)q + (\beta\tilde{z} + \tilde{\eta}) + \tilde{H}(t) \\ &= E(y_0|s,q) + (\beta\tilde{z} + \tilde{\eta}) + \tilde{H}(t). \end{aligned} \tag{B.3}$$

Now, assume that the market does not observe y_t, but receives a noisy signal of a worker's productivity each period that the worker is in the market. Denote the signal as $\xi_t = y_t + \nu_t$, where $\nu_t \sim N(0, \sigma_\nu^2)$ represents stochastic variation in a worker's productivity which is independent of s, q and η, and is assumed to be independent and identically distributed across periods. As in AP, since the market observes s and q, the production signal is equivalent to observing $I_t = \xi_t - E(y_t|s,q) = \beta\tilde{z} + \tilde{\eta} + \nu_t$, where I_t denotes a signal of the workers unobserved productive ability. Then, for a worker with t periods of experience, the filtration, or information structure, available to the market can be denoted as $\mathscr{F}_t = (I_0, \ldots, I_{t-1})$, which characterizes a worker's production history. Finally, let μ_t denote the difference between $\beta\tilde{z} + \tilde{\eta}$ and $E(\beta\tilde{z} + \tilde{\eta}|\mathscr{F}_t)$, and assume, as in AP, that μ_t is distributed independently of s, q and $\mathscr{F}_t$.

Now, given competition among firms, the AP model assumes that workers' log wages at experience level t equal their expected log-productivity, given s, q, and $\mathscr{F}_t$. This yields the following log-wage equation:

$$\omega_t = E(y_0|s,q) + H(t) + E(\beta\tilde{z} + \tilde{\eta}|\mathscr{F}_t) + \varepsilon_t, \tag{B.4}$$

where $H(t) = \tilde{H}(t) + \log(E(exp^{\mu_t}))$ represents additions to a worker's log-productivity after $t = 0$.

Notice that Equation (B.4) implies that $\omega_0 = E(y_0|s,q) + \varepsilon_0$ as the only information available to the market regarding ability is accounted for by s and q. Thus, using the log of the entry period wage, ω_0, it is possible to create a vector of time-invariant correlates of a worker's productive ability that are orthogonal to the information available to the market when the worker enters the labor market for the first time. That is, a worker's initial labor market wage should fully capture all of the characteristics that are observed by employers, including those that are not observed by the econometrician (i.e. q). Define z^* as the residual from a regression of z on the information available to the market when $t = 0$, specifically s and ω_0:

$$z^* = z - E^*(z|s,\omega_0), \tag{B.5}$$

where $E^*(z|s,\omega_0)$ is the linear projection of z on s and ω_0. Notice that since ω_0 is a function of s and q, applying the law of iterated expectations yields: $E(z|s,\omega_0) = E[E(z|s,q)|s,\omega_0] = E(z|s,q)$, implying that $E(z|s,\omega_0)$ is linear in s and ω_0 and thus equal to $E^*(z|s,\omega_0)$. Referring back to Equation (B.2), FG show that this implies that $z^* = \tilde{z}$, and thus allows the econometrician to directly control for $\tilde{z}$ at all experience levels $t > 0$.

To see how $\tilde{z}$ evolves with labor market experience, consider the conditional expectation function of ω_t given s, z^*, and t, for $t = 1, \ldots, T$ (omitting the first wage observation ($t = 0$) for each individual as it is used to generate the z^* variable),

$$E(\omega_t|s,z^*,t) = B_{st}s + B_{zt}z^* + H(t), \tag{B.6}$$

where s and z^* are reinterpreted as the components of s and z^* that are orthogonal to $H(t)$ for simplicity. Our primary interest lies in B_{zt} as this is the component that picks up employer learning regarding the portion of ability that is independent of observable characteristics. As z^* is orthogonal to both s and $H(t)$, we have that:

$$B_{zt} = \frac{\text{cov}(z^*,\omega_t)}{\text{var}(z^*)} = \frac{\text{cov}(z^*,E(\beta\tilde{z}+\tilde{\eta}|\mathscr{F}_t))}{\text{var}(z^*)}, \tag{B.7}$$

which is the standard least squares regression result (for simplicity, I have interpreted $\tilde{z}$ as a scalar).

Then, as $E(\beta\tilde{z}+\tilde{\eta}|\mathscr{F}_t)$ represents employers' updated beliefs regarding $\beta\tilde{z}+\tilde{\eta}$ after t periods of experience, we can express B_{zt} as

$$B_{zt} = \beta\rho_t, \tag{B.8}$$

where $\rho_t \in [0,1]$ represents the amount of information gained by employers through all of the observed output signals up to experience level t. Following AP, I formally describe ρ_t as,

$$\rho_t = \frac{\text{cov}(E(\beta\tilde{z}+\tilde{\eta}|\mathscr{F}_t),z^*)}{\text{cov}(\beta\tilde{z}+\tilde{\eta},z^*)} \approx \frac{\text{cov}(E(\tilde{z}|\mathscr{F}_t),z^*)}{\text{var}(z^*)},$$

which fully describes how the employer learning process evolves with experience in the symmetric learning model.

Put simply, this means that B_{zt} represents the log wage return to a worker's unobserved ability that has been learned by employers by experience level t. Additionally, since it can be shown that $\partial\rho_t/\partial t \geq 0$,[1] if we assume that for all experience levels $t > 0$ there exists some non-zero probability that $\partial\rho_t/\partial t \neq 0$, we have that $\lim_{t\to\infty} B_{zt} = \beta$, which implies that as a worker gains experience in the labor market, the observed return to their unobserved productive ability gradually moves toward equaling their true marginal productive ability.

While this characterization of the returns to the learning process used by employers in the labor market is not novel in and of itself, it will provide an important benchmark comparison for the model with event signaling developed below. The employer learning model developed in the next subsection will extend the learning model above to allow for the effects of a post labor market entry event which employers may choose to take as an additional signal of a worker's unobserved ability. I will show that by comparing

[1] See footnote 9 in AP.

the characterization of the experience-ability profile of a worker who has experienced such an event with that of the benchmark characterization developed above, my model can identify forms of post-labor market entry statistical discrimination that arise in the presence of asymmetric employer learning and non-constant observable characteristics.

B.2 A Model of Employer Learning With Unexpected Information

Now I turn to developing an extension of the model above that accounts for a post labor market entry event (*e.g.* returning to school) which can affect wages through the event itself (e.g. lost human capital), as well as through the signal it sends regarding the worker's unobserved productive ability.

Assume the same set up as above. Now, let d_{it} denote a labor market event that equals zero for all workers when $t = 0$, but at all experience levels $t > 0$, there exists some workers for whom $d_{it-1} = 0$ but $d_{it} = 1$. Further, for all $t > t_0$, $d_{it} = 1$, where t_0 is the experience level that worker i had attained when the event occurred. Importantly, outside employers in the market share a common belief about the distribution of a worker's ability, conditional on seeing that the worker experienced the event is such that $E(\eta_i^*|d_{it} = 1) \neq E(\eta_i^*)$. That is, under asymmetric information, employers that have not already directly employed a worker who experiences this type of event have a belief regarding the type of worker who is likely to experience the event.

Log productivity for a worker with t years of experience can then be expressed as (dropping the i subscript):

$$y_t = E(y_0|s,q) + \delta_d d_{t_0} + (\beta\tilde{z} + \tilde{\eta}) + \tilde{H}(t), \tag{B.9}$$

where $d_{t_0} = \mathbb{1}(t \geq t_0)$ is an indicator function equal to one if $t \geq t_0$. In this context, δ_d represents any direct effect that the labor market event may have on a worker's productivity (e.g. lost/gained human capital).

As before, the market does not directly observe a worker's output, but receives a noisy signal of the worker's production $\xi_t = y_t + \nu_t$, with $\nu_t \sim N(0, \sigma_\nu^2)$. Again, since the market observes s and q, the production signal is equivalent to observing $I_t = \xi_t - E(y_t|s,q) = \beta\tilde{z} + \tilde{\eta} + \nu_t$. Now, however, the market also must take the realization of d_{t_0} into account regarding its beliefs regarding the remaining portion of a worker's unobserved ability that has not yet been accounted for. In order to understand how the market takes the realization of d_{t_0} into account regarding ability, it is necessary to model how the market values the signaling aspect of d_{t_0} during the period it is realized.

For workers who experience the labor market event at experience level t_0, the market uses knowledge of d_{t_0} to form a new conditional expectation regarding any portion of $\beta\tilde{z}+\tilde{\eta}$ that remains unknown to the market, given the information available to it in period t_0.[2] The end result will be analogous to the conditional expectations described in AP based on knowledge of q and s, with the added caveat being that the market must also account for any additional signals received prior to the event (i.e $\mathscr{F}_{t_0}$). Let $\tilde{\eta}_{t_0} = \tilde{\eta} - E(\tilde{\eta}|\mathscr{F}_{t_0})$ be the remaining expectation error in prospective employers' beliefs about $\tilde{\eta}$ at t_0. Then prospective employers' conditional expectation function of $\tilde{\eta}_{t_0}$ with respect to the realization of d_{t_0} can be expressed as,

$$\begin{aligned}
\tilde{\eta}_{t_0} &= E(\tilde{\eta}_{t_0}|d_{t_0}) + \tilde{\eta}_d \\
&= E(\tilde{\eta} - E(\tilde{\eta}|\mathscr{F}_{t_0})|d_{t_0}) + \tilde{\eta}_d \\
&= (1-\rho_{t_0})E(\tilde{\eta}|d_{t_0}) + \tilde{\eta}_d,
\end{aligned} \tag{B.10}$$

where the second equality results from substituting $\tilde{\eta} - E(\tilde{\eta}|\mathscr{F}_{t_0})$ in for $\tilde{\eta}_{t_0}$, and the third equality results from replacing $E(\tilde{\eta}|\mathscr{F}_{t_0}) = \rho_{t_0}\tilde{\eta}$ which represents the degree to which prospective employers have already updated their beliefs regarding $\tilde{\eta}$ by experience level t_0; and $\tilde{\eta}_d$ is the remaining error in prospective employers' beliefs regarding a worker's ability. It can be shown that $\tilde{z}$ can be deconstructed in a similar manner.

In this context, $\rho_{t_0} \in (0,1)$ essentially represents a discount factor placed on the conditional expectation in order to account for the fact that prospective employers have already learned some information regarding a worker's ability based on output signals observed over the $t_0 - 1$ experience levels prior to the event. Given that $\partial\rho_{t_0}/\partial t_0 \geq 0$, the weight placed on the signaling value of the event is weakly decreasing as the experience prior to the event increases.[3] For simplicity, I assume that prospective employers' conditional expectation functions of $\tilde{\eta}$ and $\tilde{z}$ with respect to d_{t_0} are linear,

$$\begin{aligned}
\tilde{\eta}_{t_0} &= (1-\rho_{t_0})\delta_\eta d_{t_0} + \tilde{\eta}_d \\
\tilde{z}_{t_0} &= (1-\rho_{t_0})\delta_z d_{t_0} + \tilde{z}_d.
\end{aligned} \tag{B.11}$$

Now, returning to the production signals that prospective employers receive each period, for all $t > t_0$, the conditional expectation of y_t is no longer solely with respect to s and q, rather it must not take into account d_{t_0}

[2] I am assuming here that the timing of the realization of d_{t_0} occurs prior to output at experience level t_0, and thus prior to the market receiving the production signal t_0. Under this timing structure, the information available to the market regarding past production signals is $(I_o,\ldots,I_{t_0-1}) = \mathscr{F}_{t_0}$.

[3] Note that this is the ρ_t derived in the Subsection B.1, evaluated at $t = t_0$. Consequently, ρ_{t_0} can be formally expressed as

$$\rho_{t_0} = \frac{\text{cov}(E(\beta\tilde{z}+\tilde{\eta}|\mathscr{F}_{t_0}),\tilde{z})}{\text{cov}(\beta\tilde{z}+\tilde{\eta},\tilde{z})},$$

and thus $\partial\rho_{t_0}/\partial t_0 \geq 0$ follows by construction.

and the weight that prospective employers place on its ability to proxy for ability at experience level t_0. Thus, the signal received each period by prospective employers is akin to observing $I_t^d = \xi_t - E(y_t|s,q,\mathscr{F}_{t_0},d_{t_0} = 1) = \beta\tilde{z}_d + \tilde{\eta}_d + \nu_t$. As $\mathscr{F}_{t_0}$ is factored into the interpretation of I_t^d, the information set that prospective employers now use to update their beliefs about $\beta\tilde{z}_d + \tilde{\eta}_d$ can be denoted as $\mathscr{F}_{t,t_0}^d = (I_{t_0}^d,\ldots,I_{t-1}^d)$, which only takes into account production signals received following production at experience level t_0. Finally, let μ_t^d denote the difference between $\beta\tilde{z}_d + \tilde{\eta}_d$ and $E(\beta\tilde{z}_d + \tilde{\eta}_d|\mathscr{F}_t^d)$, which I assume is distributed independently of s, q, d_{t_0}, $\mathscr{F}_{t_0}$, and $\mathscr{F}_{t,t_0}^d$.

Given this, for any $t \geq t_0$, it is possible to re-write Equation (B.9) as:

$$\begin{aligned} y_t &= E(y_0|s,q) + \delta_d d_{t_0} + \overbrace{E(\beta\tilde{z}+\tilde{\eta}|\mathscr{F}_{t_0})}^{\substack{\text{Expected Ability}\\ \text{Given } \mathscr{F}_{t_0}}} + \overbrace{E(\beta\tilde{z}+\tilde{\eta}|d_{t_0})}^{\substack{\text{Expected Ability}\\ \text{Given } d_{t_0}}}\overbrace{(1-\rho_{t_0})}^{\substack{\text{Weighted}\\ \text{by } \mathscr{F}_{t_0}}} + \tilde{H}(t) + \overbrace{(\beta\tilde{z}_d+\tilde{\eta}_d)}^{\substack{\text{Remaining Error}\\ \text{Given } \mathscr{F}_{t_0} \text{ and } d_{t_0}}} \\ &= \overbrace{E(y_{t_0}|s,q,d_{t_0},\mathscr{F}_{t_0})}^{\substack{\text{Expected log-productivity}\\ \text{Given } \mathscr{F}_{t_0} \text{ and } d_{t_0}}} + \overbrace{\tilde{H}(t,t_0)}^{\substack{\text{Post-Event}\\ \text{Exp. Profile}}} + \overbrace{(\beta\tilde{z}_d+\tilde{\eta}_d)}^{\substack{\text{Remaining Error}\\ \text{Given } \mathscr{F}_{t_0} \text{ and } d_{t_0}}} \end{aligned} \tag{B.12}$$

where $\tilde{H}(t,t_0) = \tilde{H}(t) - E(\tilde{H}(t_0)|s,q,d_{t_0},\mathscr{F}_{t_0})$ represents the post-event experience profile of productivity for a worker who experiences the event in period t_0.[4,5] This yields the log-wage process given by:

$$\omega_t = E(y_{t_0}|s,q,d_{t_0},\mathscr{F}_{t_0}) + H(t,t_0) + E(\beta\tilde{z}_d + \tilde{\eta}_d|\mathscr{F}_{t,t_0}^d) + \varepsilon_t, \tag{B.13}$$

where $H(t,t_0) = \tilde{H}(t,t_0) + \log(E(exp^{\mu_t^d}))$. It is straightforward to show that when $t < t_0$, Equation (B.13) is equal to Equation (B.4), and thus a worker's log-wage should experience a discrete change at $t = t_0$, and this change should persist with additional post-event experience.

Now, define z^* as before and consider the conditional expectation function when $t = 1,\ldots,t_0,\ldots,T$,

$$E(\omega_t|s,z^*,d_{t_0},t) = B_{st}s + B_{zt}z^* + B_{dt}d_{t_0} + H(t), \tag{B.14}$$

where s and z^* are reinterpreted as components of each variable that are orthogonal to $H(t)$, conditional on d_t. Primary interest lies in B_{zt} and B_{dt}, as these components will pick up employer learning regarding the portion of ability that is independent of observable characteristics (q,s), as well as show any possible statistical discrimination that arises due to the the market observing the event d_{t_0}. It should be noted that when $t < t_0$, this conditional expectation function is the same as in Equation (B.6), and thus the result from

[4] The decomposition of the expected ability terms in the first line arises from the fact that $\tilde{\eta} = \tilde{\eta}_{t_0} + E(\tilde{\eta}|\mathscr{F}_{t_0})$ and $\tilde{z} = \tilde{z}_{t_0} + E(\tilde{z}|\mathscr{F}_{t_0})$, and then applying the result from Equation (B.10).

[5] Notice that when $t < t_0$, we have that $d_{t_0} = 0$, which implies that $E(\beta\tilde{z}+\tilde{\eta}|d_{t_0}) = 0$, and $E(\beta\tilde{z}+\tilde{\eta}|\mathscr{F}_{t_0}) = E(\beta\tilde{z}+\tilde{\eta}|\mathscr{F}_t)$, which follows from the fact that, for all $t < t_0$, $\mathscr{F}_t \subset \mathscr{F}_{t_0}$, and thus the tower property of conditional expectations implies that $E[E(\beta\tilde{z}+\tilde{\eta}|\mathscr{F}_{t_0})|\mathscr{F}_t] = E(\beta\tilde{z}+\tilde{\eta}|\mathscr{F}_t)$. Given this, when $t < t_0$, Equation (B.12) is equal to Equation (B.3).

that subsection holds. As such, most of the analysis that follows will focus on the case where $t \geq t_0$. To ease notation, I follow AP in using Φ in place of the $\text{cov}(\cdot,\cdot)/\text{var}(\cdot)$ terms below. The standard least squares omitted variable bias formula can then be expressed as,

$$\begin{aligned} B_{zt} &= \Phi_{zt_0} + \Phi^d_{zt} \\ B_{dt} &= \Phi_{dt_0} + \Phi^d_{dt} + \Phi_{qd} + H^d(t,t_0), \end{aligned} \tag{B.15}$$

where Φ_{qd} represents the coefficient from a regression of $\tilde{\alpha}q$ on d_{t_0}, Φ_{zt_0} is the coefficient from the regression of $E(\beta\tilde{z}+\tilde{\eta}|\mathscr{F}_{t_0})$ on z^*, Φ_{dt_0} is the coefficient from the regression of $\tilde{\delta}d_{t_0}$ on d_{it},[6] Φ^d_{zt} and Φ^d_{dt} represent the coefficients from the regressions of $E(\beta\tilde{z}_d+\tilde{\eta}_d|\mathscr{F}^d_{t,t_0})$ on z^* and d_{it} respectively, and $H^d(t,t_0)$ represents the direct effect of d_{t_0} on a worker's experience profile of productivity.[7]

When $t = t_0$, both Φ^d_{zt} and Φ^d_{dt} are zero by construction due to the fact that $E(\beta\tilde{z}_d+\tilde{\eta}_d|\mathscr{F}^d_{t,t_0}) = 0$, as discussed above. Additionally, it is easy to show that $\Phi_{zt_0} = \beta\rho_{t_0}$, as this follows from the analogous least squares regression result derived in Subsection B.1, evaluated at $t = t_0$, while $\Phi_{dt_0} = \delta_d + (1-\rho_{t_0})(\delta_\eta + \beta\delta_z)$, which comes mechanically from the regression of $\tilde{\delta}$ on d_{t_0}. Denote the combined return to the unobserved ability portion of B_{zt} and B_{dt} at experience level $t = t_0$ that is accounted for by $\tilde{z}$ as

$$R_{zt_0} = \underbrace{\beta\rho_{t_0}}_{\substack{\text{Return to Output} \\ \text{Signals at } t=t_0}} + \overbrace{(1-\rho_{t_0})}^{\substack{\text{Weighted by Output} \\ \text{Signals Through } t_0}} \underbrace{\beta\delta_z}_{\substack{\text{Unconditional} \\ \text{Event Signal}}} \tag{B.16}$$

where the second term comes directly from the relationship between $\tilde{z}$ and d_{t_0} derived in Equation (B.11). This illustrates that a prospective employer's overall assessment of a worker's unobserved ability following some labor market event is a convex combination of what the employer's assessment would have been in the absence of the event and the employer's beliefs regarding ability conditional on the labor market event.

Now, when $t > t_0$, Φ^d_{zt} and Φ^d_{dt} must be accounted for in interpreting the returns to unobserved ability following the labor market event, and how the effect of the event itself evolves with experience. To account for these variables, I deconstruct these terms following AP as

$$\begin{aligned} \Phi^d_{zt} &= \rho^d_{t'}\Phi^d_z \\ \Phi^d_{dt} &= \rho^d_{t'}\Phi^d_d, \end{aligned}$$

[6]Where $\tilde{\delta} = \delta_d + (1-\rho_{t_0})(\delta_\eta + \beta\delta_z)$.

[7]Implicit in this setup is the assumption that the manner in which employers learn about the remaining portion of a worker's unobserved ability is orthogonal to the post event experience profile of productivity $H^d(t,t_0)$.

where Φ_z^d and Φ_d^d represent the coefficients from the regression of $\beta\tilde{z}_d + \tilde{\eta}_d$ on z^* and d_{t_0} respectively, and

$$\rho_{t'}^d = \frac{\text{cov}(E(\beta\tilde{z}_d + \tilde{\eta}_d|\mathscr{F}_{t'}^d), z^*)}{\text{cov}(\beta\tilde{z}_d + \tilde{\eta}_d, z^*)} \in [0,1],$$

which represents the degree to which prospective employers have learned about their remaining expectation error regarding a worker's ability between experience levels t_0 and t, with $t' = t - t_0$. It can then be shown that $\Phi_z^d = \beta(1-\rho_{t_0})$ and $\Phi_d^d = -(1-\rho_{t_0})(\delta_\eta + \beta\delta_z)$, where the second expression comes directly from the derivation in Equation (B.11).

From this, I can formally express B_{zt}^d as,

$$B_{zt} = B_{zt_0} + b_{zt'} + b_{z(t',t_0)} \tag{B.17}$$

where $B_{zt_0} = \beta\rho_{t_0}$ represents what prospective employers had learned about a worker's ability through experience level t_0, while $b_{zt'} = \beta\rho_{t'}^d$ represents the degree to which employer's have updated their beliefs regarding $\tilde{z}_d$ as a function of post-event experience, and $b_{z(t',t_0)} = -\beta\rho_{t_0}\rho_{t'}^d$ represents the relationship between pre- and post-event learning.[8] This representation is the basis for the following proposition.

Proposition 5 Under the assumptions of the model developed above, if an event d is such that prospective employers share the common belief that $E(\beta\tilde{z}+\tilde{\eta}|d=1) \neq E(\beta\tilde{z}+\tilde{\eta})$, then for sufficiently small t_0 and t,

$$\frac{\partial B_{zt}|d=1}{\partial t'} > \frac{\partial B_{zt}|d=0}{\partial t'}.$$

Simply put, Proposition 5 arises from the fact that when $t > t_0$, B_{zt} accounts for the discrete shift in learning that takes place by prospective employers as they shift from learning about their initial expectation error regarding a worker's unobserved ability when the worker first entered the labor market $(\beta\tilde{z}+\tilde{\eta})$, to learning about their updated expectation error regarding a worker's ability following the labor market event $(\beta\tilde{z}_d + \tilde{\eta}_d)$.[9] Put another way, B_{zt} picks up the difference in the post-event belief updating processes for workers who experienced the event $((1-\rho_{t_0})\rho_{t'}^d)$ relative to those who did not $(\rho_t - \rho_{t_0})$ that is due to the difference in prospective employers' expectation errors regarding unobserved ability before and after the labor market event.

[8]This last term accounts for the fact that the amount of post-event learning that can take place is directly related to how much learning already had taken place prior to the labor market event - if a significant amount of learning had already occurred prior to the event, then there is a limited amount of learning that can still take place after the event.

[9]More formally, Proposition 5 is a result of the fact that $\text{var}(\beta\tilde{z}+\tilde{\eta}|d=1) < \text{var}(\beta\tilde{z}+\tilde{\eta})$ as the event must reduce the variance of the distribution of unobserved ability in order for employers to use it as a meaningful signal. If the event does not reduce the variance in the distribution of unobserved ability for those workers, the market will ignore the signal.

Moving on to B_{dt}, at any experience level $t \geq t_0$, employers are updating the weight that they apply to the d_{t_0} signal as they gain additional output signals regarding the worker's true ability. Formally, it can be shown that this can be expressed as

$$B_{dt} = \underbrace{\delta_d}_{\substack{\text{True Effect}\\ \text{of Event}}} + \underbrace{b_{zd} - B_{zt}\delta_z}_{\substack{\text{Signal Value of Event}\\ \text{at Expierence Level } t}} + \underbrace{\Phi^d_{\tilde{\eta}(t,t_0)} + \Phi_{qd} + H^d(t,t_0)}_{\substack{\text{Additinal Determinants of the}\\ \text{Effect of the Event That are } \perp z^*.}} \tag{B.18}$$

where $b_{zd} = \beta\delta_z$ represents the unconditional effect of the event signal and $B_{zt}\delta_z$ is the return to the ability-experience learning profile defined in Equation (B.17) weighted by the event signal (δ_z).

Taken together with Equation (B.17), this shows the channels through which pre- and post-event experience will reduce the weight placed on the d_{t_0} signal in any give period. This leads to the following proposition.

Proposition 6 Under the assumptions of the model above,

$$\frac{\partial B_{dt}}{\partial t'} = -\delta_z\left(\frac{\partial b_{zt'}}{\partial t'} + \frac{\partial b_{z(t',t_0)}}{\partial t'}\right) + \frac{\partial H^d(t,t_0)}{\partial t'} = -\delta_z\frac{\partial B_{zt}}{\partial t'} + \frac{\partial H^d(t,t_0)}{\partial t'}$$
$$\frac{\partial B_{dt}}{\partial t_0} = -\delta_z\left(\frac{\partial B_{zt_0}}{\partial t_0} + \frac{\partial b_{z(t',t_0)}}{\partial t_0}\right) + \frac{\partial H^d(t,t_0)}{\partial t_0} = -\delta_z\frac{\partial B_{zt}}{\partial t_0} + \frac{\partial H^d(t,t_0)}{\partial t_0},$$

and the portion of B_{dt} that accounts for the signaling nature of some event d decreases to zero in the limit of both t_0 and t.

Intuitively, as t_0 or t' increase, prospective employers learn more about the worker's true unobserved ability, decreasing the need to rely on the event signal as a proxy for information on the worker's ability to the point where no information is gained by observing the event.

The above can also be seen by looking at how the combined return to unobserved ability from Equation (B.17) and Equation (B.18) evolves with both pre- and post-event experience. To see this, for any experience level $t \geq t_0$, denote the combined return to the unobserved ability portion of B_{zt} and B_{dt} that is accounted for by $\tilde{z}$ as

$$R_{zt} = \underbrace{B_{zt_0}}_{\substack{\text{Return to Pre-Event}\\ \text{Experience at } t_0}}\overbrace{(1-\delta_z)}^{\substack{\text{Relative to the}\\ \text{Event Signal}}} + \underbrace{b_{z(t',t_0)}}_{\substack{\text{Return to Function of}\\ \text{Pre and Post Event Exp}}}\overbrace{(1-\delta_z)}^{\substack{\text{Relative to the}\\ \text{Event Signal}}} + \underbrace{b^d_{zt'}}_{\substack{\text{Return to Post-Event}\\ \text{Experience at } t'}}\overbrace{(1-\delta_z)}^{\substack{\text{Relative to the}\\ \text{Event Signal}}} + \underbrace{b_{zd}}_{\substack{\text{Unconditional}\\ \text{Event Signal}}} \tag{B.19}$$

This equation illustrates how a prospective employer's overall assessment of a worker's unobserved ability following some labor market event evolves with both pre- and post-event experience. Essentially this illustrates that as additional output signals are received by the employer and ρ_{t_0} or $\rho^d_{t'}$ increase, not only will the worker benefit (or be hurt) by the increased emphasis toward β, but they will also benefit (or be hurt) by decreased emphasis on the signaling value of the event. This forms the basis of the next proposition.

Proposition 7 Under the assumptions of the model developed above, if an event d is such that $E(\eta^*|d = 1) < E(\eta^*)$, then for $t \geq t_0$ the following are true,

$$\frac{\partial R_{zt}}{\partial t'} > \frac{\partial B_{zt}|d=0}{\partial t'} > 0, \quad \frac{\partial R_{zt}}{\partial t_0} > \frac{\partial B_{zt}|d=0}{\partial t_0} > 0, \quad and \quad \frac{\partial^2 R_{zt}}{\partial t' \partial t_0} < \frac{\partial^2 B_{zt}}{\partial t' \partial t_0} < 0.$$

On the other hand, if $E(\beta\tilde{z} + \tilde{\eta}|d = 1) > 0$ and $\delta_z > 1$, then for $t \geq t_0$ the following are true,

$$\frac{\partial R_{zt}}{\partial t'} < 0 < \frac{\partial B_{zt}|d=0}{\partial t'}, \quad \frac{\partial R_{zt}}{\partial t_0} < 0 < \frac{\partial B_{zt}|d=0}{\partial t_0}, \quad and \quad \frac{\partial^2 R_{zt}}{\partial t' \partial t_0} > 0 > \frac{\partial^2 B_{zt}}{\partial t' \partial t_0}.$$

This shows that if the ability signal is such that δ_z is negative (positive), the total return to the ability-experience profile for some variable z^* that is observed by the econometrician but not prospective employers after an event ($\frac{\partial R_{zt}}{\partial t}$) will be larger (smaller) than the relative return to ability-experience profile ($\frac{\partial B_{zt}}{\partial t}$), and this difference will be directly proportional to the relative size of the event signal δ_z. The general implications of this proposition are illustrated in Figure 2.2.

At this point it is worth observing that R_{zt}^d represents the true coefficient on z^* from a wage regression that fails to include d_{t_0}, and thus represents what is being estimated in employer learning models that assume symmetric information (as in the baseline case in the previous subsection). As such, the importance of the implications of Proposition 7 for interpreting employer learning models that assume symmetric employer learning is quite clear: if information regarding a worker's true ability is not symmetric, and prospective employers take observed "events" as signals regarding a worker's true ability, then the estimated effects from the symmetric learning models will be biased, and this bias will persist even when the event itself is accounted for if how the event changes the ability-experience profile following the event is not.

Taken together, B_{zt} and R_{zt}^d should allow this model to identify specific aspects of employer learning in the labor market that the benchmark models could not by treating these signaling "events" (*e.g.* returning to school) as a form of statistical discrimination. That is, the model should be able to identify if prospective employers respond to the signaling nature of some event, given the information available at the time by looking at how the pre- and post-event ability-experience profiles evolve.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abadie, Alberto and Matias D. Cattaneo (2018). "Econometric Methods for Program Evaluation." *Annual Review of Economics* 10, 465–503.

Acharya, Viral, Ramin P. Baghai, and Krishnamurthy V. Subramanian (2013). "Labor Laws and Innovation." *The Journal of Law and Economics* 56 (4), 997–1037.

Addison, John T., Liwen Chen, and Orgul D. Ozturk (2020). "Occupational Skill Mismatch: Differences by Gender and Cohort." *ILR Review* 73 (3), 730–767.

Altonji, Joseph G., Prashant Bharadwaj, and Fabian Lange (2012a). "Changes in the Characteristics of American Youth: Implications for Adult Outcomes." *Journal of Labor Economics* 30 (4), 783–828.

Altonji, Joseph G., Prashant Bharadwaj, and Fabian Lange (2012b). "Changes in the Characteristics of American Youth: Implications for Adult Outcomes." *Journal of Labor Economics* 30 (4), 783–828.

Altonji, Joseph G., Lisa B. Kahn, and Jamin D. Speer (2016). "Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success." *Journal of Labor Economics* 34 (S1).

Altonji, Joseph G. and Charles R. Pierret (1998). "Employer Learning and the Signalling Value of Education." In: *Internal Labour Markets, Incentives and Employment*. Ed. by Isao Ohashi and Toshiaki TachibanakiEditors. Palgrave Macmillan, London, 159–195.

Altonji, Joseph G. and Charles R. Pierret (2001). "Employer Learning and Statistical Discrimination." *The Quarterly Journal of Economics* 116 (1), 313–350.

Angelov, Nikolay, Per Johansson, and Erica Lindahl (2016). "Parenthood and the Gender Gap in Pay." *Journal of Labor Economics* 34 (3), 545–579.

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom (2016). "College Attrition and the Dynamics of Information Revelation." NBER Working Paper No. 22325, National Bureau of Economic Research, Cambridge, M.A..

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo (2010). "Beyond Signaling and Human Capital: Education and the Revelation of Ability." *American Economic Journal: Applied Economics* 2 (4), 76–104.

Arellano-Bover, Jaime (2020). "Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size." IZA Discussion Paper No. 12969.

Arellano-Bover, Jaime (2021). "The Effect of Labor Market Conditions at Entry on Workers' Long-Term Skills." *Review of Economics and Statistics* 2. Forthcoming.

Ashworth, Jared, V. Joseph Hotz, Arnaud Maurel, and Tyler Ransom (2020). "Changes across Cohorts in Wage Returns to Schooling and Early Work Experiences." *Journal of Labor Economics*. Forthcoming.

ASVAB Career Exploration Program (2011). *The ASVAB Career Exploration Program: Theoretical and Technical Underpinnings of the Revised Skill Composites and OCCU-Find*. Tech. rep.

Bacolod, Marigee P. and Bernardo S. Blum (2010). "Two Sides of the Same Coin: Residual Inequality and the Gender Gap." *Journal of Human Resources* 45 (1), 197–242.

Barach, Moshe A. and John J. Horton (2020). "How Do Employers Use Compensation History?: Evidence From a Field Experiment." NBER Working Paper No. 26627, National Bureau of Economic Research, Cambridge, M.A..

Barnette, Justin, Kennedy Odongo, and C. Lockwood Reynolds (2021). "Changes Over Time in the Cost of Job Loss for Young Men and Women." *BE Journal of Economic Analysis and Policy* 21 (1), 335–378.

Barron, John M., Mark C. Berger, and Dan A. Black (2006). "Selective Counteroffers." *Journal of Labor Economics* 24 (3), 385–409.

Bates, Michael (2019). "Public and Private Employer Learning: Evidence from the Adoption of Teacher Value-Added." *Journal of Labor Economics*. Forthcoming.

Beaudry, Paul, David A. Green, and Benjamin M. Sand (2016). "The Great Reversal in the Demand for Skill and Cognitive Tasks." *Journal of Labor Economics* 34 (S1).

Bedard, Kelly (2001). "Human Capital versus Signaling Models: University Access and High School Dropouts." *Journal of Political Economy* 109 (4).

Blau, Francine D. and Lawrence M. Kahn (2017). "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* 55 (3), 789–865.

Blinder, Alan S. (1973). "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources* 8 (4), 436–455.

Burdett, Kenneth, Carlos Carillo-Tudela, and Melvyn Coles (2020). "The Cost of Job Loss." *The Review of Economic Studies* 87 (4), 1757–1798.

Carrington, William J. and Bruce Fallick (2017). "Why Do Earnings Fall with Job Displacement?" *Industrial Relations* 56 (4), 688–722.

Castex, Gonzalo and Evgenia Kogan Dechter (2014). "The Changing Roles of Education and Ability in Wage Determination." *Journal of Labor Economics* 32 (4), 685–710.

Chung, YoonKyung, Barbara Downs, Danielle H. Sandler, and Robert Sienkiewicz (2017). "The Parental Gender Earnings Gap in the United States." *Center for Economic Studies Discussion Paper 17-68*.

Cohn, Alain, Michel André Maréchal, Frédéric Schneider, and Roberto A. Weber (2021). "Frequent Job Changes Can Signal Poor Work Attitude and Reduce Employability." *Journal of the European Economic Association* 19 (1), 475–508.

Cortes, Patricia and Jessica Pan (2020). "Children and the Remaining Gender Gaps in the Labor Market."

NBER Working Paper 27980.

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

Davis, Steven and John Haltiwanger (2014). "Labor Market Fluidity and Economic Performance," NBER Working Paper No. 20479, National Bureau of Economic Research, Cambridge, M.A..

Deming, David J. (2017). "The Growing Importance of Social Skills in the Labor Market." *Quarterly Journal of Economics* 132 (4), 1593–1640.

Devaro, Jed and Michael Waldman (2012). "The Signaling Role of Promotions: Further Theory and Empirical Evidence." *Journal of Labor Economics* 30 (1), 91–147.

Doiron, Denise J. (1995). "Lay-Offs As Signals: The Canadian Evidence." *The Canadian Journal of Economics* 28 (4a), 899–913.

Fallick, Bruce (1993). "The Industrial Mobility of Displaced Workers." *Journal of Labor Economics* 11 (2), 302–323.

Fan, Xiadong and Jed DeVaro (2020). "Job Hopping and Adverse Selection in the Labor Market." *The Journal of Law, Economics, and Organization* 36 (1), 84–138.

Farber, Henry (2015). "Job Loss in the Great Recession and its Aftermath: U.S. Evidence from the Displaced Workers Survey." NBER Working Paper No. 21216, National Bureau of Economic Research, Cambridge, M.A..

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

Farber, Henry S. and Robert Gibbons (1996). "Learning and Wage Dynamics." *Quarterly Journal of Economics* 111 (4), 1007–1047.

Farber, Henry and Robert Gibbons (1996). "Learning and wage dynamics." *The Quarterly Journal of Economics* 111 (4), 1007–1047.

Forsythe, Eliza (2019). "Careers within Firms: Occupational Mobility Over the Lifecycle." *LABOUR* 33 (3), 241–277.

Forsythe, Eliza (2020). "Occupational Job Ladders and Displaced Workers." Unpublished Working Paper.

Frazis, Harley (2002). "Human capital, signaling, and the pattern of returns to education." *Oxford Economic Papers* 54 (2), 298–320.

Fuji, Mayu, Kousuke Shiraishi, and Noriyuki Takayama (2018). "The Effects of Early Job Separation on Later Life Outcomes." *Journal of the Japanese and International Economies* 48, 68–84.

Gibbons, Robert and Lawrence F. Katz (1991). "Layoffs and Lemons." *Journal of Labor Economics* 9 (4), 351–380.

Goldin, Claudia (2014). “A Grand Gender Convergence: Its Last Chapter.” *American Economic Review* 104 (4), 1091–1119.

Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko (2006). “The Homecoming of American College Women: The Reversal of the College Gender Gap.” *Journal of Economic Perspectives* 20 (4), 133–156.

Greenwald, Bruce C. (1986). “Adverse Selection in the Labour Market.” *The Review of Economic Studies* 53 (3), 325–347.

Grund, Christian (1999). “Stigma Effects of Layoffs?: Evidence From German Micro-data.” *Economics Letters* 64 (2), 241–247.

Guvenen, Faith, Burhan Kuruscu, Satoshi Tanaka, and David Wiczer (2020). “Multidimensional Skill Mismatch.” *American Economic Journal: Macroeconomics* 12 (1), 210–244.

Hall, Robert E. and Alan B. Krueger (2012). “Evidence on the Incidence of Wage Posting, Wage Bargaining, and On-the-Job Search.” *American Economic Journal: Macroeconomics* 4 (4), 56–67.

Haltiwanger, John, Henry R. Hyatt, Lisa B. Kahn, and Erika Mcentarfer (2018). “Cyclical Job Ladders by Firm Size and Firm Wage.” *American Economic Journal: Macroeconomics* 10 (2), 52–85.

Haltiwanger, John, Stefano Scarpetta, and Helena Schweiger (2014). “Cross country differences in job reallocation: The role of industry, firm size and regulations.” *Labour Economics* 26, 11–25.

Harris, Milton and Bengt Holmstrom (1982). “A Theory of Wage Dynamics.” *The Review of Economic Studies* 49 (3), 315–333.

Heckman, James J., Jora Stixrud, and Sergio Urzua (2006). “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior.” *Journal of Labor Economics* 24 (3), 411–482.

Hershbein, Brad J. (2012). “Graduating High School in a Recession: Work, Education, and Home Production.” *The B.E. Journal of Economic Analysis & Policy* 12 (1).

Hsin, Amy and Christina Felfe (2014). “When Does Time Matter? Maternal Employment, Children’s Time with Parents, and Child Development.” *Demography* 51, 1867–1894.

Hu, Luojia and Christopher Taber (2011). “Displacement, Asymmetric Information, and Heterogeneous Human Capital.” *Journal of Labor Economics* 29 (1), 113–152.

Hussey, Andrew (2012). “Human capital augmentation versus the signaling value of MBA education.” *Economics of Education Review* 31 (4), 442–451.

Jacobson, Louis S, Robert J LaLonde, and Daniel G. Sullivan (1993). “Earnings Losses of Displaced Workers.” *The American Economic Review* 83 (4), 685–709.

Jarosch, Gregor (2021). "Searching for Job Security and the Consequences of Job Loss." NBER Working Paper No. 28481, National Bureau of Economic Research, Cambridge, M.A..

Kahn, Lisa B (2010). "The long-term labor market consequences of graduating from college in a bad economy." *Labour Economics* 17 (2), 303–316.

Kahn, Lisa B (2013). "Asymmetric Information between Employers." *American Economic Journal: Applied Economics* 5 (4), 165–205.

Keane, Michael P. and Kenneth I. Wolpin (1997). "The Career Decisions of Young Men." *Journal of Political Economy* 105 (3), 473–522.

Kletzer, Lori G. (1998). "Job Displacement." *Journal of Economic Perspectives* 12 (1), 115–136.

Kletzer, Lori G. and Robert W. Fairlie (2003). "The Long-Term Costs of Job Displacement for Young Adult Workers." *ILR Review* 56 (4), 682–698.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller (2019). "Child Penalties Across Countries: Evidence and Explanations." *AEA Papers and Proceedings* 109, 122–126.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard (2019). "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* 11 (4), 181–209.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard (2021). "Does Biology Drive Child Penalties? Evidence from Biological and Adoptive Families." *American Economic Review: Insights* 3 (2), 183–198.

Kosovich, Stephen M. (2010). "The Value of Layoffs and Labor Market Conditions as Signals of Worker Quality." *The B.E. Journal of Economic Analysis & Policy* 10 (1).

Kostol, Andreas R (2017). "Mismatch and the Consequence of Job Loss." *Preliminary Working Paper*.

Krashinsky, Harry (2002). "Evidence on Adverse Selection and Establishment Size in the Labor Market." *Industrial and Labor Relations Review* 56 (1), 84–96.

Krolikowski, Pawel (2017). "Job Ladders and Earnings of Displaced Workers." *American Economic Journal: Macroeconomics* 9 (2), 1–31.

Lachowska, Marta, Alexandre Mas, and Stephen Woodbury (2020). "Sources of Displaced Workers' Long-Term Earnings Losses." *American Economic Review* 110 (10), 3231–3266.

Lange, Fabian (2007). "The Speed of Employer Learning." *Journal of Labor Economics* 25 (1), 1–35.

Lengermann, Paul A. and Lars Vilhuber (2002). *Abandoning the Sinking Ship: The Composition of Worker Flows Prior to Displacement*. Technical Paper TP-2002-11. LEHD, U.S. Census Bureau.

Light, Audrey (1995). "Hazard model estimates of the decision to reenroll in school." *Labour Economics* 2 (4), 381–406.

Light, Audrey and Kathleen McGarry (1998). "Job Change Patterns and the Wages of Young Men." *The Review of Economics and Statistics* 80, 276–286.

Light, Audrey and Andrew McGee (2015a). "Does employer learning vary by schooling attainment? The answer depends on how career start dates are defined." *Labour Economics* 32, 57–66.

Light, Audrey and Andrew McGee (2015b). "Employer Learning and the "Importance" of Skills." *Journal of Human Resources* 50 (1), 72–107.

Lise, Jeremy and Fabien Postel-Vinay (2020). "Multidimensional Skills, Sorting, and Human Capital Accumulation." *American Economic Review* 110 (8), 2328–2376.

Liu, Kai (2019). "Wage Risk and the Value of Job Mobility in Early Employment Careers." *Journal of Labor Economics* 37 (1), 139–185.

Liu, Kai, Kjell G. Salvanes, and Erik Ø. Sørensen (2016). "Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession." *European Economic Review* 84, 3–17.

Mansour, Hani (2012). "Does Employer Learning Vary by Occupation?" *Journal of Labor Economics* 30 (2), 415–444.

Michaud, Amanda M. (2018). "A Quantitative Theory of Information, Worker Flows, and Wage Dispersion." *American Economic Journal: Macroeconomics* 10 (2), 154–183.

Moscarini, Giuseppe and Fabien Postel-Vinay (2012). "The Contribution of Large and Small Employers to Job Creation in Times of High and Low Unemployment." *American Economic Review* 102 (6), 2509–2539.

Mroz, Thomas A. and Timothy H. Savage (2006). "The Long-Term Effects of Youth Unemployment." *The Journal of Human Resources* 41 (2), 259–293.

Mukoyama, Toshihiko and Sophie Osotimehin (2019). "Barriers to Reallocation and Economic Growth: the Effects of Firing Costs." *American Economic Journal: Macroeconomics*. Forthcoming.

Murphy, Kevin M. and Finis Welch (1990). "Empirical Age-Earnings Profiles." *Journal of Labor Economics* 8 (2), 202–229.

Nakamura, Emi (2008). "Layoffs and Lemons Over the Business Cycle." *Economics Letters* 99 (1), 55–58.

Neal, Derek (1995). "Industry-Specific Human Capital: Evidence from Displaced Workers." *Journal of Labor Economics* 13 (4), 653–677.

Neal, Derek (1999). "The Complexity of Job Mobility among Young Men." *Journal of Labor Economics* 17 (2), 237–261.

Neal, Derek and William Johnson (1996). "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* 104 (5), 869–895.

Nedelkoska, Ljubica, Frank Neffke, and Simon Wiederhold (2015). "Skill Mismatch and the Costs of Job Displacement." *Preliminary Working Paper*.

Neumark, David (2002). "Youth Labor Markets in the United States: Shopping Around vs. Staying Put." *Review of Economics and Statistics* 84 (3), 462–482.

Oaxaca, Ronald (1973). "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14 (3), 793–709.

Oyer, Paul and Scott Schaefer (2000). "Layoffs and Litigation." *The RAND Journal of Economics* 31 (2), 345–358.

Oyer, Paul and Scott Schaefer (2011). "Personnel Economics: Hiring and Incentives." In: *Handbook of Labor Economics*. Ed. by David Card and Orley AshenfelterEditors. Vol. 4B. Great Britian: North Holland. Chap. 20, 1769–1823.

Petre, Melinda (2018). "Are Employers Omniscient? Employer Learning About Cognitive and Noncognitive Skills." *Industrial Relations* 57 (3), 323–360.

Pinkston, Joshua C. (2003). "Screening discrimination and the determinants of wages." *Labour Economics* 10 (6), 643–658.

Pinkston, Joshua C. (2006). "A Test of Screening Discrimination with Employer Learning." *ILR Review* 59 (2), 267–284.

Pinkston, Joshua C. (2009). "A Model of Asymmetric Employer Learning with Testable Implications." *Review of Economic Studies* 76 (1), 367–394.

Postel-Vinay, Fabien and Jean-Marc Robin (2004). "To match or not to match?: Optimal wage policy with endogenous worker search intensity." *Review of Economic Dynamics* 7 (2), 297–330.

Postel-Vinay, Fabien and Hélène Turon (2013). "The Impact of Firing Restrictions on Labour Market Equilibrium in the Presence of On-the-job Search." *The Economic Journal* 124 (575), 31–61.

Riley, John (2001). "Silver Signals: Twenty-Five Years of Screening and Signaling." *Journal of Economic Literature* 39 (2), 432–478.

Rothstein, Jesse (2020). "The Lost Generation? Labor Market Outcomes for Post Great Recession Entrants." NBER Working Paper No. 27516, National Bureau of Economic Research, Cambridge, M.A..

Schönberg, Uta (2007a). "Testing for Asymmetric Employer Learning." *Journal of Labor Economics* 25 (4), 651–691.

Schönberg, Uta (2007b). "Testing for Asymmetric Employer Learning." *Journal of Labor Economics* 25 (1), 651–691.

Schwandt, Hannes and Till von Wachter (2019). "Unlucky Cohorts: Estimating the Long-Term Effects of Entering the Labor Market in a Recession in Large Cross-Sectional Data Sets." *Journal of Labor*

Economics 37 (S1), S161–S198.

Schwerdt, Cuido (2011). “Labor Turnover Before Plant Closure: ’Leaving the Sinking Ship’ vs. ’Captain Throwing Ballast Overboard’.” *Labour Economics* 18 (1), 93–101.

Segall, Daniel O. (1997). “Equating the CAT-ASVAB.” in: *Computerized Adaptive Testing: From Inquiry to Operation*. Ed. by William A. Sands, Brian K. Waters, and James R. McBride. American Psychological Association. Washington, DC.

Seim, David (2019). “On the Incidence and Effects of Job Displacement: Evidence from Sweden.” *Labour Economics* 57, 131–145.

Sieppi, Antti and Jaakko Pehkonen (2019). “Parenthood and Gender Inequality: Population-Based Evidence on the Child Penalty in Finland.” *Economics Letters* 182, 5–9.

Sims, William H. and Catherine Hiatt (2001). *Marine Corps Selection and Classification*. Tech. rep. Center for Naval Analysis.

Song, Younghwan (2007). “Recall bias in the displaced workers survey: Are layoffs really lemons?” *Labour Economics* 14 (3), 335–345.

Speer, Jamin D. (2017a). “Pre-Market Skills, Occupational Choice, and Career Progression.” *Journal of Human Resources* 52 (1), 187–246.

Speer, Jamin D. (2017b). “Pre-Market Skills, Occupational Choice, and Career Progression.” *Journal of Human Resources* 52 (1), 187–246.

Spence, Michael (1973). “Job Market Signaling.” *The Quarterly Journal of Economics* 87 (3), 355–374.

Spence, Michael (2002). “Signaling in Retrospect and the Informational Structure of Markets.” *American Economic Review* 92 (3), 434–459.

Stevens, Ann Huff (1997). “Persistent Effects of Job Displacement: The Importance of Multiple Job Losses.” *Journal of Labor Economics* 15 (1, Part 1), 165–188.

Stinebrickner, Ralph and Todd Stinebrickner (2012). “Learning about Academic Ability and the College Dropout Decision.” *Journal of Labor Economics* 30 (4), 707–748.

Swinkels, Peroen M. (1999). “Education Signalling with Preemptive Offers.” *The Review of Economic Studies* 66 (4), 949–970.

Topel, Robert H. (1990). “Specific Capital and Unemployment: Measuring the Costs and Consequences of Job Loss.” *Carnegie-Rochester Conference Series on Public Policy* 33, 181–214.

Topel, Robert H. and Michael P. Ward (1992). “Job Mobility and the Careers of Young Men.” *The Quarterly Journal of Economics* 107 (2), 439–479.

Tyler, John H., Richard J. Murnane, and John B. Willett (2000). “Estimating the Labor Market Signaling

Value of the GED." *The Quarterly Journal of Economics* 115 (2), 431–468.

U.S. Bureau of Labor Statistics (2020). *Employee Benefits Survey*.

VanderBerg, Bryce (2021a). "Employer Learning and Dynamic Statistical Discrimination." Preliminary Working Paper.

VanderBerg, Bryce (2021b). "The Signaling Role of Early Career Job Loss." Unpublished Working Paper.

von Wachter, Till and Stefan Bender (2006). "In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers' Careers." *American Economic Review* 96 (5), 1679–1705.

Waldman, Michael (1984). "Job assignment, signalling, and efficiency." *Rand Journal of Economics* 15 (2), 255–267.

Waldman, Michael (2012). "Theory and Evidence in Internal Labor Markets." In: *The Handbook of Organizational Economics*. Ed. by Robert S. Gibbons and John RobertsEditors. Princeton University Press, Princeton, NJ, 520–574.

Welsh, John R., Susan K. Kucinkas, and Linda T. Curran (1990). *Armed Services Vocational Battery (ASVAB): Integrative Review of Validity Studies*. Tech. rep. 90-22. Brooks Air Force Base, TX: Air Force Systems Command.

Zhang, Ye (2007). "Employer Learning Under Asymmetric Information: The Role of Job Mobility." *SSRN Electronic Journal*.