ARC Centre of Excellence in Population Ageing Research

Working Paper 2015/23

Does Job Hopping Help or Hinder Careers? The Signaling Role of Work History

Xiaodong Fan[1] and Jed De Varo[2]

[1]Research Fellow, Centre of Excellence in Population Ageing Research (CEPAR), Australian School of Business, University of New South Wales email: x.fan@unsw.edu.au
[2]Departments of Management and Economics, California Sate University, USA email: jed.devaro@csueastbay.edu

Does Job Hopping Help or Hinder Careers? The Signaling Role of Work History*

Xiaodong Fan† Jed DeVaro‡

April 24, 2015

Abstract

Using NLSY data, we show that job hopping is associated with lower wages for college graduates (but not high school graduates), controlling for ability, labor market experience, and current job tenure. The effect is most pronounced for job tenures less than one year, strongest for early-career workers, and mitigated when job hopping severs matches that were formed during economic downturns. A model of employer learning is proposed in which prior job history signals worker ability. We analyze it both for symmetric and asymmetric employer learning, showing that the equilibrium under asymmetric learning is consistent with the evidence for college graduates.

KEYWORDS: job hopping, asymmetric employer learning, work history

JEL classification: D82, J31, J63

*We would like to thank John Kennan, Rasmus Lentz, Christopher Taber, and Michael Waldman for invaluable discussion and inspiration. We are also grateful to Joseph Altonji, Hugh Cassidy, Hanming Fang, Michael Keane, Hodaka Morita, Uta Schönberg, Alan Woodland, and seminar participants at Wisconsin and UNSW for helpful comments.

†ARC Centre of Excellence in Population Ageing Research (CEPAR), University of New South Wales, Sydney, NSW 2052, Australia. Email: x.fan@unsw.edu.au.

‡Departments of Management and Economics, College of Business and Economics, California State University, East Bay, 25800 Carlos Bee Boulevard, Hayward, CA 94542-3066, USA. Email: jed.devaro@csueastbay.edu.

"...*Decide what you feel you're worth and, once you're satisfied with your salary level, try to stay in one place for a while.*" *(Jessica Kleiman, "8 Pros and Cons of Job Hopping", Forbes, 8/06/2012.)*

1 Introduction

The typical career spans multiple employers, and switching employers after a short job tenure is particularly common for early-career workers.[1] But whether such "job hopping" helps or hinders careers is an open question. One perspective is that job hopping signals negative attributes (e.g. disloyalty, impatience, a short attention span, low productivity, and a high likelihood of future turnover) that scare off prospective employers. Evidence from employer surveys and case studies indeed suggests that job hopping creates a stigma that can hurt careers.[2] An alternative perspective is that experience in multiple jobs spanning a number of employers provides skills that the labor market values and that can enhance careers. Such experience broadens information and skill sets and may create versatile, jacks-of-all-trades who are well equipped for success, particularly in managerial and executive roles. That argument is at the heart of Lazear's entrepreneurship model and is consistent with his empirical evidence (from a sample of alumni of Stanford University's MBA program) that experience in many different roles is the most important determinant of becoming an entrepreneur (Lazear, 2005). The following conclusion from a 2014 survey by CareerBuilder, the largest online employment website in the US, reflects both perspectives on how employers view job hoppers: "*43 percent of employers will not consider a candidate having short tenures with several employers, while 53 percent of employers said job-hoppers tend to have a wide range of expertise and can adapt quickly (51 percent).*"[3]

An important starting point for understanding the effect of job hopping on careers is to examine the effect of a worker's job turnover history on his or her current wage. The first goal of this study is to answer that question using data from the National Longitudinal Survey of Youth 1979 (NLSY79). For workers with only a high school degree, we find little or no effect of work history on the current wage. But we find that college graduates suffer a job hopping penalty or, in other words, enjoy a reward for "loyalty" to the employer. Their current wages are negatively

[1] A 2012 Bureau of Labor Statistics (BLS) study based on National Longitudinal Survey of Youth 1979 (NLSY79) data finds that the average American born between 1957 and 1964 held 11.3 jobs from age 18 to 46, and that nearly half of those jobs were held from ages 18 to 24. See also Topel and Ward (1992), which finds in the Longitudinal Employee-Employer Data (LEED) between 1957 and 1972 that a typical worker has 7 jobs during the first 10 years of labor market experience.

[2] For example, a 2012 survey by Bullhorn, a recruiting software company, finds that 39 percent of 1500 recruiters and hiring managers say that "*the single biggest obstacle for an unemployed candidate in regaining employment is having a history of 'hopping jobs,' or leaving a company before one year of tenure.*" And in case studies of 6 Chicago organizations, Bills (1990) finds a) that many employers use job history information in making hiring or promotion decisions, and b) that the most important component of those decisions is job hopping, which can be detrimental for both middle and low status positions. A similar result is found in a local restaurant industry (Bills, 1999).

[3] An analysis of the advantages and disadvantages of job hopping at the level of a particular industry is offered in Fallick et al. (2006), which argues that a benefit of job hopping in California's Silicon Valley is the rapid reallocation of resources towards firms with superior innovations, though a cost of that job hopping is human capital externalities that reduce the incentives to invest in new knowledge.

correlated with their numbers of previous employers, after controlling for ability, labor market experience, and current job tenure. The negative effect of job hopping on current wages is especially pronounced for job tenures less than one year. We also find that college graduates' wage penalty from job hopping is larger for early-career workers and that it is mitigated when workers enter the labor market during an economic downturn.

The second goal of the study is to provide a theoretical framework for interpreting the preceding evidence and, in doing so, to connect our empirical findings for high school and college graduates to the growing literature on employer learning (about worker ability) in the labor market. The new feature of our analysis is that a worker's prior job history (which may include job hopping) signals the worker's ability to employers in the labor market. In the employer learning literature, a worker's current wage reflects all the productivity-related information that is held by the worker's current employer and by other prospective employers. That literature has two branches. One assumes symmetric learning, meaning that all employers in the labor market learn about a worker's ability at the same rate.[4] Another assumes asymmetric learning, meaning that a worker's current employer (henceforth the incumbent employer) has private information about that worker's ability, based on having witnessed that worker perform on the job; competing, prospective employers lack that information.[5] Although both perspectives are supported in the empirical literature, for higher-skilled, more highly-educated workers, recent evidence points to the importance of asymmetric learning (e.g. Schönberg (2007) and Kim and Usui (2014) find evidence of symmetric learning for high school graduates and asymmetric learning for college graduates, and DeVaro and Waldman (2012) and Cassidy et al. (2015) find evidence of a signaling role of promotions only for workers with at least a college degree).

In the preceding literature, the source of employer learning is usually either direct observation of worker productivity on-the-job, or indirect observation via the *most recent* tenure or job assignment (i.e. a promotion, or a lack of promotion). But as the documentation in our opening paragraph suggests, job histories are readily observable via résumés, and employers base hiring and wage decisions on those entire histories. Thus, it is natural that a worker's job history, which might involve job hopping, would signal information about ability and future productivity.

Due to the complexity arising from asymmetric learning, most models that make that informational assumption have only two periods, which by assumption rules out the role of work history.

[4]Representative studies include Harris and Holmstrom (1982), Farber and Gibbons (1996), Gibbons and Waldman (1999, 2006), Altonji and Pierret (2001), Lange (2007), Antonovics and Golan (2012), and DeVaro and Morita (2013).

[5]In the asymmetric learning literature, the incumbent employer typically observes the worker's ability or productivity, perhaps with noise, whereas competing employers only observe signals, such as a worker's voluntary or involuntary turnover (Greenwald, 1986; Gibbons and Katz, 1991; Schönberg, 2007; Pinkston, 2009; Hu and Taber, 2011; Zhang, 2011; Kahn, 2013), the firm's job assignment or promotion decisions (Waldman, 1984, 1990, 1996; DeVaro and Waldman, 2012), training levels (Acemoglu and Pischke, 1998), or wages (Golan, 2009).

Empirical support for asymmetric learning is found in Gibbons and Katz (1991), Pinkston (2009), Hu and Taber (2011), and Kahn (2013), and also in Schönberg (2007) and Kim and Usui (2014) for college graduates. A subset of the asymmetric learning literature has focused on the role of job assignments (and particularly promotions) as signals of worker ability, following Waldman (1984), Bernhardt (1995), Owan (2004), and the related analysis of Milgrom and Oster (1987). Evidence from a recent empirical literature supports such asymmetric learning via promotions (e.g. DeVaro and Waldman 2012, Bognanno and Melero 2014, Cassidy et al. 2015, and DeVaro et al. 2015).

That is, when prospective employers consider hiring decisions at the start of the second period, workers have (at most) experience with only one prior employer and have not had a chance to job hop. We prove the existence of a general equilibrium in a three-period asymmetric learning model where some workers hop jobs. Workers in the model make turnover decisions at the end of the first and second periods—either remaining with the incumbent employer or quitting to join a competing employer—in consideration of the signaling effect of those decisions for the future. Competing employers observe all prior turnover decisions and make inferences about workers' abilities, accordingly. Therefore, workers make turnover decisions and self-select (or self-label) based on their own productivities. High-productivity workers switch jobs less often than low-productivity workers, thereby signaling to competing employers that they are not "lemons" (Akerlof, 1970), which in turn results in longer tenures with current and previous employers. Thus, if employer learning is asymmetric, competing employers interpret both current and previous job tenures as signals of worker ability and make wage offers accordingly. In contrast, if learning is symmetric, prior job tenures have no effect on current wages, after controlling for ability, tenure with the current employer, and labor market experience. That observation can explain our differential empirical results for high school and college graduates and forms the basis for a new empirical test of asymmetric versus symmetric learning in the labor market.

In the model, productivity is jointly determined by ability and a worker-firm-specific match quality component that generates turnover and allows an analysis of the welfare loss associated with inefficient turnover.[6] When employer learning is symmetric (or when ability is publicly observed), workers with poor match qualities switch firms to seek better matches, and these separations are independent of ability. However, when abilities are not observable and employer learning is asymmetric, high-ability workers with poor match quality might find it optimal to stay with their current employers (to avoid being seen as lemons in the eyes of the market) while low-ability workers with good match quality might switch firms to take advantage of "the benefit of the doubt." Such inefficient turnover would not arise in a world of publicly observable ability or symmetric learning. The basic logic of the result in the asymmetric learning case can be described as the successive application of Greenwald's (1986) adverse selection-turnover argument when there is a possibility of multiple turnover decisions in the worker's career.

A main result of our theoretical analysis is that when work history signals worker ability, the symmetric (asymmetric) learning model is consistent with our empirical evidence for high school (college) graduates. A further implication, under asymmetric learning, is that the standard approach of including labor market experience and current job tenure (but not prior work history) on the right-hand side of a wage regression yields biased estimates of the returns to tenure and labor market experience. Additionally, anything that should make prospective employers sympathetic to a job hopping event (e.g. a quit that severs a match formed during an economic downturn, since such a separation is relatively more likely to have arisen from a poor match than from low

[6] "Inefficient turnover" refers to the termination of high-quality matches or the persistence of low-quality matches, neither of which would occur in a symmetric learning model.

ability) should mitigate the wage penalty under asymmetric learning.

Several studies of asymmetric employer learning examine the signaling role of job histories and are therefore related to ours, though none of them yields our testable implications or explains why college graduates face a wage penalty for job hopping. Greenwald (1986) develops a similar three-period asymmetric employer learning model, but that study is differentiated from ours in that Greenwald does not prove the existence of the equilibrium and does not derive (or empirically test) our testable implications. Another point of differentiation is that incorporating a worker-firm-specific match quality allows for analysis of the potential welfare loss caused by inefficient turnover, whereas in Greenwald (1986) turnover does not result in welfare losses.

In another theoretical analysis, Golan (2009) studies a three-period model in which abilities are revealed in wages through bargaining, and Zhang (2011) studies a three-period model in a partial equilibrium analysis, ignoring workers' response to the signaling effect of work history. DeVaro and Waldman (2012) analyze a two-period model and informally discuss an extension to three periods, but because of firm-specific human capital there is no equilibrium turnover in that model. Similarly, Bernhardt (1995) analyzes a three-period model in which workers are differentiated by education levels and promotions signal worker ability, but because the equilibrium exhibits no turnover, there are no results concerning job hopping.

Evidence that the relevance of job hopping and short job tenures in the labor market has increased is suggested in the analysis of Farber (2008) using CPS data from 1973 to 2006, where it is shown that long-term employment relationships have become much less common for men in the private sector and that the same worker group has seen an increase in "churning" (defined as the proportion of workers in jobs with less than a year of tenure). The 2012 survey evidence from Bullhorn, cited in footnote 2, suggests that such "churning" sends a negative signal to employers about prospective hires, consistent with our empirical evidence and theoretical arguments.

The empirical evidence we report relates to a literature on mover-stayer models, and in particular to Munasinghe and Sigman (2004), which adds a job mobility measure to the right-hand side of a wage regression estimated on NLSY data.[7] In that study, the key "job mobility" regressor is the ratio of the number of prior employers to labor market experience. The numerator of that ratio is the key regressor in our analysis. That study finds a negative coefficient on the key mobility regressor, suggesting prior mobility is associated with lower current wages. Apart from using the numerator rather than the ratio as the key regressor, our empirical analysis differs in three significant ways. First, we show that the negative relationship masks a stark difference between education groups, i.e. it exists for college graduates but not for high school graduates. Second, we identify a critical threshold of job tenure (namely one year) under which the job hopping wage penalty for college graduates is particularly sharp. Third, we find evidence that the job hopping wage penalty for college graduates is mitigated when workers enter the labor market during an economic downturn. The theoretical framework used in Munasinghe and Sigman (2004) to ra-

[7]See that study for a discussion of the mover-stayer literature. It should be understood that, unlike the theoretical framework we develop, the mover-stayer framework focuses only on whether the worker stays or moves at the most recent opportunity and does not refer to the entire prior job history.

tionalize their empirical findings also differs markedly from ours. That analysis uses a partial equilibrium model in which firms play no role (and therefore there is no learning about worker ability), the wage distribution is exogenous, and there are two periods. Ours is an equilibrium model in which firms learn about workers' abilities via observation of their prior job histories, and the model has three periods to allow for different patterns of job hopping. Career patterns and wages are endogenously determined by the wage offers of employers and the job hopping decisions of workers.

2 Does Job Hopping Help or Hinder Careers?

Using the 1979 cohort of the National Longitudinal Survey of Youth data (NLSY79), we investigate how work history affects the current hourly log-wage. Consider the following regression model:

$$y_{it} = X_{it}'\beta + WH_{it}'\gamma + \varepsilon_{it}, \tag{1}$$

where y_{it} is the logarithm of worker i's current hourly wage at time t, the vector X_{it} denotes the set of both time-invariant individual characteristics and time-varying individual current employment status (including labor market experience and tenure with the current firm), ε_{it} is an i.i.d disturbance term with mean zero, and the vector WH_{it} contains measures of work history, which may include job hopping.

By "work history" we mean one's entire labor market experience prior to joining the current employer. For reasons of data availability and measurement difficulties, "work history" regressors are excluded from the typical wage regressions appearing in the literature. The usual approach is to control only for overall experience (i.e. time in the labor market, regardless of how that time is allocated across different employers) and current job tenure (i.e. time with the current employer). A complete set of measures describing the entire work history and capturing the timing of each separation is impractical in most cases.[8] Therefore, we use a summary statistic (the number of prior employers) to proxy for the full work history. We also conduct robustness checks using an alternative summary statistic for work history (namely, the length of the shortest prior tenure). Most of our analysis focuses on continuous work histories.

2.1 The NLSY79 Data

The NLSY79 is a nationally representative sample of men and women who were 14-22 years old in 1979 and who were interviewed annually up to 1994 and biennially thereafter. We focus on males in the private sector from 1979 to 2008. The data include each worker's entire (weekly) work history from the point of labor market entry. We restrict the analysis to periods after each individual is observed entering the labor market for the first time, which is defined as "primarily"

[8] No currently available data set is large enough to accommodate even a 10-period analysis. A work history of N periods of consecutive employment would generate 2^{N-1} different turnover paths, given that a worker starts with a particular employer in the first period and can switch employers in any subsequent period.

working for at least two consecutive years (Schönberg, 2007).[9] We choose one job with the longest aggregated hours worked for each month for each individual, identifying the employers for this job and linking them in continuous years to form the job turnover history. To ensure a complete work history, we eliminate all subsequent periods after the first gap of two or more years absent from the labor market for each individual.

For each year in which a worker is observed, we keep track of three summary statistics which partially capture the entire labor market history, the first two of which summarize the worker's current labor market status, and the third one of which summarizes prior jobs: (1) the length of current job tenure, i.e. the total weeks worked for the current employer; (2) experience, i.e. total weeks worked for all employers since first entering the labor market; (3) the number of previous employers one had prior to joining the current firm. The third measure captures the likelihood of being a job hopper, because the more frequently one "hops" jobs, the larger the number of previous jobs becomes. We do not consider unemployment spells; our focus is on the effect of prior job tenures on the current wage.

The wage measure is the hourly rate of pay. Since wage information is only collected at the annual level, we keep one observation for each job in each year (from the monthly employment history data) and construct a work history by year. Hourly wages below \$1 or above \$500 (in 2004 dollars) are set to missing and are dropped from the regression, though tenure is still included in the work history construction.

The NLSY contains an ability measure, namely the Armed Forces Qualifications Test (AFQT) score, that we use as a control as in Neal and Johnson (1996), Altonji and Pierret (2001), Schönberg (2007) and Pinkston (2009). It is standardized to a mean of zero and a standard deviation of one after filtering out age-specific fixed effects among males. The time-invariant individual characteristics relevant to wage determination include dummy variables for race and US residency status at the age of 14. The time-varying controls include years of schooling and its interaction with experience, the interaction between AFQT score and experience (Altonji and Pierret, 2001), year dummies and their interactions with years of schooling, a dummy variable for urban residency, a dummy variable indicating the interview year (Schönberg, 2007), and dummy variables for occupation and industry codes at the one-digit level.[10]

We consider college graduates and high school graduates separately. College graduates are defined to have sixteen or more years of schooling, and high school graduates are those with at most twelve years of schooling.[11] Our final analysis sample consists of 923 college graduates with

[9] An individual is said to be primarily working if they work at least 1000 hours per year without changing the highest grade of education completed. Of 6403 males, 1848 are never observed to meet these criteria and are therefore excluded.

[10] As in Schönberg (2007), we distinguish seven occupations: professional, technical, and kindred; managers and administrators; sales, clerical, and kindred; craftsmen and kindred; operatives; laborers; and service. We distinguish twelve industries: mining; construction; manufacturing; transportation, communications, and other public utilities; wholesale trade; retail trade; finance, insurance, and real estate; business and repair services; personnel services; entertainment and recreation services; professional and related services; and public administration.

[11] Unreported results for those with some college (defined as 13 to 15 years) are available upon request; there are 992 individuals with 20,253 observations in this group. Results for that group are similar to (but smaller in magnitude than) those for college graduates.

18,544 observations and 2949 high school graduates with 61,058 observations.[12] Table 1 displays descriptive statistics for that sample. The average real hourly wage is $24.84 in 2004 dollars for college graduates and $14.11 for high school graduates. On average, each individual has held 6.6 jobs if he has a college degree or 6.5 jobs if he has a high school degree only, between 1979 and 2008. Most individuals were US residents at age 14, and many (i.e. 85% of college graduates and 76% of high school graduates) live in urban areas.

2.2 Current Wages and Prior Job Hopping

Table 2 displays the results of log-wage regressions for college graduates (column 1) and high school graduates (column 4), which control for ability (via AFQT score), current tenure, and experience. Of primary interest is the marginal effect of the first covariate, the number of previous jobs. Column 1 reveals that one additional employer in one's work history is associated with a 2.5 percent decrease in the current wage for workers with at least a college degree. In contrast, for workers with only a high school degree the corresponding effect in column 4 is statistically insignificant and small in magnitude. These results suggest a wage penalty from job hopping for college graduates, but not high school graduates, even after controlling for current job tenure, labor market experience, and ability.[13,14] We also find, as in Altonji and Pierret (2001) and Arcidiacono et al. (2010) using the same data but without controlling for work history, that the marginal effect of the AFQT score is positive and increasing in labor market experience for high school graduates.

Since job hopping is more common among young workers (Topel and Ward, 1992), we investigate whether the job hopping wage penalty for college graduates differs by stage of career. We first partition the sample into a "rookies" group (i.e. those with five or fewer years in the labor market) and a "veterans" group (i.e. those with more than five years in the labor market). The main result concerning a job-hopping wage penalty for college graduates (column 1 of table 2) also holds for rookies (column 2) and veterans (column 3). The result from column 4 that there is no job-hopping wage penalty for high school graduates is also true for rookies (column 5) and veterans (column 6).

[12]The analysis sample excludes 246 individuals with missing AFQT scores, eight for whom US residency status is missing, and 531 who entered the labor market before age 14 or after age 28 (both inclusive). We drop individuals with missing education levels, those with missing occupation or industry codes, those who are farmers or workers in the agriculture, forestry and fisheries industry, and those in the military, thereby eliminating another 46 individuals.

[13]Our restriction of the sample to periods after each individual is first observed working at least 1000 hours per year for at least two consecutive years without changing schooling levels ensures that we choose a labor market entry point after which each worker is fully engaged. But this rule might omit early-career periods (e.g. the transition from school to work) during which workers often sample jobs (Topel and Ward, 1992; Neal, 1999). Employers might view job hopping during these periods differently and perhaps with greater sympathy. As a robustness check to allow for this possibility, we relaxed the selection criterion by defining labor market entry as working at least 100 hours per year for at least one year. For this broadened sample the result for high school graduates is unchanged (i.e. -0.004 and statistically insignificant, as in column 4 of table 2); the result for college graduates is qualitatively unchanged (i.e. negative and statistically significant at the one percent level) though somewhat smaller in magnitude (i.e. -0.012 rather than -0.025 as in column 1 of table 2).

[14]Ariga et al. (2012) find that job hopping is associated with wage growth for migrants with less than college education in China. Only jobs that last six months or more are counted in that analysis.

The main results from table 2 are robust to using an alternative measure of job hopping in the work history, namely the length of the shortest prior job tenure. If a worker engages in frequent job hopping, the shortest prior tenure is likely to be quite short. Therefore, a positive coefficient on this measure indicates that current wages are higher when the shortest prior tenure is longer, which is consistent with a job-hopping wage penalty. As seen in table 3, the coefficient of this alternative measure is positive and statistically significant for college graduates but not high school graduates, consistent with our results from table 2 based on the number of prior employers. A one-year increase in the shortest prior job tenure for college graduates is associated with a 1.1% increase in the current wage, whereas for workers with only a high school degree the corresponding effect is small and statistically insignificant.[15] Thus, using the number of prior employers or the alternative measure of work history, the bottom line is that job hopping is associated with a wage penalty for college graduates but not high school graduates, and this negative effect is stronger for early-career workers among college graduates. As a further robustness check, we control for a non-cognitive skill measure available in the NLSY (a summation of the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale) and find that the results are largely unchanged.[16]

The 2012 survey by Bullhorn, cited in footnote 2, suggests that leaving a job before one year of tenure sends a very negative signal. To investigate whether there are "critical cutoffs" for job tenure (such as one year) that a worker must clear to avoid a wage penalty in subsequent jobs, we define a dummy variable equaling one if tenure is less than a given threshold value, and zero otherwise. We then estimate the log-wage regressions replacing the previous measures of work history that appeared in WH_{it} with a new variable, i.e. the number of previous job tenures (not including the present one) that fall below the threshold. We consider a sequence of different threshold values (all in months), i.e. 3, 6, 9, 12, ..., 36.

Figure 1 plots the results. The left panel (for college graduates), plots coefficients of the number of previous tenures shorter than the threshold value indicated on the horizontal axis, along with the 90% confidence interval. The right panel repeats the plots for those workers with only a high school degree. Two patterns are of interest. First, the left panel shows that tenures shorter than one year (exclusive) are associated with considerably larger (in magnitude) wage penalties for college graduates than tenures longer than a year. This result is consistent with the quote from the 2012 Bullhorn survey in footnote 2. Second, for high school graduates (right panel), the effects of the numbers of previous tenures shorter than certain threshold value are all near zero and statistically insignificant at the 10% level for all threshold values ranging from three months to three years.

[15]Recent labor market entrants or workers with long work histories that are concentrated only in one or two firms pose problems for regressions that include the shortest prior job tenure (along with current job tenure) on the right-hand side. But excluding workers from the sample who never had turnover leaves our results substantively unchanged. Our results are also robust to various specification changes following the literature on asymmetric employer learning. In particular, including interactions between current tenure and schooling, as well as interactions between current tenure and AFQT scores, as in Schönberg (2007), leaves our results substantively unchanged. Similarly, simultaneously including the continuous employment spell and its interactions with schooling and the AFQT scores, as in Pinkston (2009), leaves our results qualitatively unchanged.

[16]Some studies find that non-cognitive skills, as well as the cognitive skills that would be captured by AFQT scores, affect labor market outcomes (e.g., Heckman et al., 2006).

A potentially important omitted variable in our regression models is the macroeconomic condition that prevails at the time when college graduates enter the labor market.[17] To explore this possibility, we include in the regression model underlying table 2 an interaction between the national college unemployment rate in the labor market entry year and the work history. Letting u_{t_0} denote this unemployment rate, where t_0 is the year of labor market entry, the augmented regression model is:

$$\ln w_{it} = X_{it}'\beta + WH_{it}'\gamma + \varphi u_{t_0} + u_{t_0} \cdot WH_{it}'\xi + \epsilon_{it}, \tag{2}$$

Estimation results for rookie college graduates (i.e. those in their first five years in the labor market) appear in column 1 of table 4. The estimate of φ is negative (i.e. -0.048) and statistically significant, as first documented by Kahn (2010) using the same data but without controlling for work history. Including the unemployment control leaves our main result from table 2 preserved. That is, the marginal effect of the number of previous jobs is negative and statistically significant, though it is mitigated (given that the estimated ξ is positive and statistically significant) when macroeconomic conditions at the time of labor market entry are poor. For a worker graduating from college and entering the labor market in a good year with a low unemployment rate of 2%, the detrimental effect of one additional previous employer is -6.4%.[18] When the unemployment rate increases by one percentage point to 3%, the detrimental effect of one additional previous employer becomes -2.9%. Further investigation reveals that such mitigation is concave, as shown in column 2 when we include the quadratic term of the unemployment rate and its interaction with the number of previous jobs.[19] These results are consistent with employers having some degree of sympathy for job hopping when workers have the misfortune of entering the labor market during a bad economy.

Results for veteran college graduates (columns 3 and 4) reveal a wage penalty from job hopping that is smaller in magnitude than for rookies and, moreover, that is not mitigated by poor macroeconomic conditions at the time of labor market entry (i.e. the estimated ξ, is negative and statistically insignificant).[20] The difference in results for rookie versus veteran college graduates suggests that the effect of the state of the economy at the point of labor market entry wears off with experience. For veterans, the effect of macroeconomic conditions at the point of labor market entry may fade out over time because the economy may have significantly improved.[21] Another interpretation is that bad economic conditions may be more damaging (in the sense of poor match qualities) for new labor market entrants than for those with a longer record of experience. That

[17] Kahn (2010) finds in the NLSY79 data that the national (or state) college graduate unemployment rate at the time of college graduation has a large, significantly negative and persistent effect on college graduates' labor market outcomes, and Oreopoulos et al. (2012) find similar patterns in Canadian data.

[18] $-0.134 + 0.035 \times 2 = -0.064$.

[19] This implies that it is unlikely that the total effect of job hopping will become significantly positive. The estimates in table 4 imply that the total effect of the number of previous jobs is always negative.

[20] However, the overall effect of the number of previous jobs is still negative and statistically significant.

[21] Altonji et al. (2014) find that the effects of macroeconomic conditions at the point of labor market entry fade out after the first seven years.

would justify employers assigning lower wage penalties to the former group than to the latter when matches are formed in an economic downturn.

All of the preceding results are weaker for high school graduates than for college graduates, as seen in columns 5-8. For rookies, the pattern of signs of the key coefficients matches that of column 1 for rookie college graduates, though all of the magnitudes are considerably smaller. Moreover, the overall effect of the number of previous jobs in column 5 evaluated at the average unemployment rate is statistically insignificant at the 10% level. Recall also that the dependent variable is in logs, so the marginal effects in column 5 (versus those in column 1) represent smaller percentage changes starting from a smaller base. For veterans with a high school degree only (column 7), none of the aforementioned effects is statistically significant. Therefore, for high school graduates the job hopping stigma is at most minimal for rookies and is insignificant for veterans when we control for macroeconomic conditions in the labor market entry year.

A limitation of wage regressions estimated on NLSY data should be acknowledged, namely that there may be omitted variables bias due to unobservable firm or match-specific characteristics (e.g. the NLSY data lack detailed information on specific firms or worker-firm specific matches).[22] Detailed worker-firm matched data may offer opportunities to account for such unobserved heterogeneity, though such data sets typically lack measures of worker ability (both cognitive and non-cognitive) that are particular strengths of the NLSY.

To summarize, college graduates suffer a wage penalty from job hopping, and it is particularly sharp for job tenures of less than one year. That wage penalty is larger for rookies (i.e. those in their first five years of labor market experience) than for veterans. For high school graduates, there is little or no wage penalty for job hopping. These results are robust to various econometric specifications. Furthermore, the wage penalty that rookie college graduates incur from job hopping is mitigated when workers enter the labor market during an economic downturn.

3 A Model of Employer Learning from Job Histories

Our goal in developing a theoretical model is to provide an organizing framework for interpreting and better understanding the preceding new empirical facts. The model we develop—in which employers learn about worker ability by observing job histories—is a natural analytical framework to use for thinking about job hopping. The signaling role of one's current labor market status, including current turnover and job assignments, has been a central focus of the theoretical literature on promotion and wage dynamics since Waldman (1984) and Greenwald (1986). The key point in that literature is that potential employers can typically observe an applicant's current tenure or position by reading the applicant's CV and then using that information to draw inferences about the applicant's ability. Our point is that employers are capable of reading more than one line of the CV (i.e. the most recent job tenure or assignment) and that they pay attention to the entire work history. The employer learning framework is helpful for interpreting the evidence in

[22] See Buchinsky et al. (2010) for an overview of these issues.

the preceding section in that the forthcoming theoretical predictions map closely to the preceding empirical facts.[23]

Time is discrete, and the model has three periods. At least three periods are needed to study the effects of job histories. The term "job history" (or "work history") refers to the turnover history prior to the current job. For example, if someone works for a different firm in each of the three periods, then during the third period the "job history" has two different firms prior to the current job. We develop the analysis for both symmetric and asymmetric learning, given the relevance of both perspectives in the prior literature. We show that both perspectives are helpful for interpreting our empirical findings. The main result is that under symmetric learning, a worker's current wage is independent of the work history, whereas under asymmetric learning, there is a wage penalty when the work history exhibits job hopping. Thus, recalling our main empirical finding in table 2, the symmetric learning case is more applicable to high school graduates, and the asymmetric learning case is more applicable to college graduates.

3.1 Symmetric Learning

Consider a three-period model with workers, their employers (i.e. the incumbent employers), and a market of competing firms.[24] There is a measure one of workers in the labor market. One worker can only work for one employer at a time. Firms and workers are risk neutral and do not discount the future.[25] Each worker is endowed with (and privately observes) a time-invariant general ability, θ, that cannot be credibly revealed or signaled to prospective employers at the start of the game. At the end of the first period, all employers observe a worker's ability and stay informed during all subsequent periods.

There is free entry of firms and perfect competition for workers. Firms can only make spot wage contracts for one period and cannot commit to future wage contracts, which implies an equilibrium condition that expected profit must be zero for any new worker hired in any period. At the beginning of the first period, potential firms, who do not observe worker ability, enter the labor market and simultaneously post wages. After observing posted wages, each worker accepts an offer. We are interested in the symmetric equilibrium in which firms with the same information make identical wage offers. In that case, all firms simultaneously post identical first-period wage offers of V_1, which are high enough to ensure zero expected profit. Each worker then randomly chooses an offer, and the market clears instantaneously.

For a matched worker-firm pair, the first-period output, y_1, is determined by both ability and

[23] However, even if that had not been so (i.e. if we had found empirically a job hopping *reward* rather than a penalty), we believe the forthcoming analysis would have been of interest. In that case our focus would have been on trying to ferret out the reason for the incompatibility between the empirical evidence and natural theoretical frameworks, including the forthcoming one. It is surely possible to write down a reasonable theoretical model of job hopping that predicts a reward rather than a penalty, but even in that case one would want to understand why the employer learning framework fails.

[24] We use "firms" and "employers" interchangeably.

[25] Discounting complicates the equilibrium computation but does not change the analysis qualitatively.

worker-firm-specific match quality, η_1, i.e. $y_1 = \theta + \eta_1$.[26] The distributions of ability—$F(\theta)$—and of match quality—$G(\eta)$—are independent and common knowledge. We assume $F(\theta)$ and $G(\eta)$ are continuous in the supports $[\theta_L, \theta_H]$ and $[\eta_L, \eta_H]$, respectively, where any of the bounds can be infinite. We assume that $F(\theta)$ is log-concave and that it is not proportional to e^{cx}, with $c > 0$.[27]

Match quality, η_1, is an experience good, which means the worker does not observe it before accepting the wage offer. At the end of the first period, each worker's ability, θ, and match quality, η_1, are revealed to all employers and workers. We assume that η_1 remains fixed for as long as the worker remains with the employer, though this is not crucial for our results.[28] Although we assume that match quality with the initial employer is observed by all employers, our analysis is unaffected if competing employers lack this information. The reason is that a new, independent match quality is drawn for each new worker-firm pair, which implies that the wage offers from competing firms are independent of the current match quality even if it is observable, as long as θ is public information.[29]

At the end of the first period, all firms simultaneously post two second-period wage offers—one for incumbent workers, $V_2(\theta, \eta_1)$, and one for other workers, $V_2^o(\theta)$—which are observed by all workers.[30] We refer to these as the incumbent offer and market offer, respectively. Our results are insensitive to allowing the incumbent firm to make a counteroffer after observing the highest market offer, since in equilibrium the latter is correctly anticipated.

Each worker decides which second-period offer to accept. If a worker accepts the incumbent offer, the previous match quality, η_1, remains throughout the second period, and second-period output is $y_2 = y_1 = \theta + \eta_1$.[31] If a worker accepts the market offer and therefore quits the incumbent firm to work for a new firm, a new match quality, η_2, is independently drawn from $G(\eta)$. The realization of η_2 is not observed by the worker or by employers until the end of the second period. The second-period output for the new matched worker-firm pair is $y_2 = \theta + \eta_2$.

At the end of the second period, two wage offers are again made, the incumbent offer, $V_3(\theta, \eta_j)$ where $j \in \{1, 2\}$, and the market offer, $V_3^o(\theta)$. Each worker chooses the highest offer. We assume no recall, i.e. a firm does not retain any information about a previous worker once a match

[26]The match quality component generates equilibrium turnover. Its importance to wage determination is well established in the literature (Johnson, 1978; Jovanovic, 1979; Neal, 1999; Nagypal, 2007). It also generates equilibrium turnover. Previous work has generated turnover by assuming exogenous separations (Greenwald, 1986) or stochastic utility shocks (Acemoglu and Pischke, 1998; Schönberg, 2007). An advantage of generating turnover via a match quality component is that it facilitates welfare analysis of the efficiency loss due to asymmetric learning.

[27]These are rather weak assumptions. Distributions that satisfy them are the normal, uniform, exponential, and extreme value, among others (Bagnoli and Bergstrom, 2005).

[28]If match quality were instead increasing in job tenure, the main difference is that the equilibrium quitting probability would decrease.

[29]A "partial" symmetric learning model, in which other employers only observe productivity but not each of its components, is beyond the scope of this study.

[30]The market offer is independent of the previous match quality, since that match quality becomes irrelevant upon separation.

[31]For simplicity, we do not include a productivity increase as in Gibbons and Katz (1991), which may be interpreted as firm-specific human capital accumulation. Including such a component would not change our analysis as long as it is not so large that it shuts down turnover completely.

breaks.[32] Assuming the worker switches jobs after the second period, a new, independent match quality, η_3, is drawn for the newly-formed pair, which is revealed to all parties at the end of the third period when all workers retire from the labor market.

The following proposition describes the equilibrium (see the Appendix for all proofs).

Proposition 1: *In the symmetric learning model, at the end of the first period, a worker stays with the incumbent firm if and only if* $\eta_1 \geq \eta^*_{s1}$, *where*

$$\eta^*_{s1} = \bar{\eta} + \frac{1}{2} E \left\{ (\eta_2 - \bar{\eta}) \cdot \mathbf{1} \left\{ \eta_2 \geq \bar{\eta} \right\} \right\}. \tag{3}$$

The market wage offer is

$$V^o_2 (\theta) = \theta + \bar{\eta} + E \left\{ (\eta_2 - \bar{\eta}) \cdot \mathbf{1} \left\{ \eta_2 \geq \bar{\eta} \right\} \right\}, \tag{4}$$

and the incumbent wage offer is

$$V_2 (\theta, \eta_1) = \min \left\{ \theta + \eta_1 + (\eta_1 - \bar{\eta}) \cdot \mathbf{1} \left\{ \eta_1 \geq \bar{\eta} \right\}, V^o_2 (\theta) \right\}. \tag{5}$$

At the end of the second period, a worker stays with the incumbent firm if and only if $\eta_j \geq \bar{\eta}$, *where* $j \in \{1, 2\}$. *The market wage offer is*

$$V^o_3 (\theta) = \theta + \bar{\eta}. \tag{6}$$

The incumbent wage offer is

$$V_3 (\theta, \eta_j) = \min \left\{ \theta + \eta_j, V^o_3 (\theta) \right\} = \theta + \min \left\{ \eta_j, \bar{\eta} \right\}. \tag{7}$$

In both the second and third periods, the accepted incumbent offer matches the market offer.

Given free entry to the labor market, the first-period wage is determined by the zero expected profit condition,

$$\begin{aligned} V_1 \;=\; & \bar{\theta} + \bar{\eta} + E \left\{ \left[\eta_1 - \bar{\eta} - E \left\{ (\eta_2 - \bar{\eta}) \cdot \mathbf{1} \left\{ \eta_2 \geq \bar{\eta} \right\} \right\} \right] \cdot \mathbf{1} \left\{ \eta_1 \geq \eta^*_{s1} \right\} \right. \\ & \left. + (\eta_1 - \bar{\eta}) \cdot \mathbf{1} \left\{ \eta_1 \geq \bar{\eta} \right\} \right\}. \end{aligned} \tag{8}$$

Proposition 1 states that under symmetric learning the worker's turnover decision is independent of ability and depends entirely on the realization of match quality. Workers with poor match qualities quit to seek better matches elsewhere, and those with good match qualities stay with their incumbent employers. Under symmetric learning there is no adverse selection in the labor market.

[32]Therefore, if a worker happens to end up with the same employer in the first and third periods (but a different one in the second period), then in the third period the employer does not recall the worker's first-period ability, and a fresh match quality is drawn. For this reason, "firm" and "job" are used interchangeably throughout the analysis. The preceding point concerning recall is irrelevant under symmetric learning but will be relevant under asymmetric learning.

3.2 Asymmetric Learning

All notation and modeling assumptions from the preceding subsection also apply in the asymmetric learning case unless otherwise noted. Under asymmetric learning, the incumbent employer privately observes θ and η_1 at the end of the first period. Competing firms do not observe any information about the employment relationship—ability, match quality, output, or wage. Let W_1 denote the first-period wage offer.[33] Compared to the preceding analysis, the new information structure changes wage offers for the second and third periods.

We consider two types of workers, namely "stayers" and "quitters". At the end of the first period, the incumbent second-period wage offer becomes $W_2(y_1)$, since the incumbent firm cares only about productivity and, therefore, offers a wage as a function of y_1 only, rather than θ or η_1. Let W_2^o denote the market wage offer. A worker quits if and only if $W_2(y_1) < W_2^o$, and the worker is assumed to stay if indifferent. Competing firms can only make inferences about ability from observed quitting behavior. Stayers and quitters are defined according to their decisions at the end of the first period and are indexed by "s" and "q" in subsequent notation. Figure 2 shows the timing.

The equilibrium of this three-period asymmetric learning model is solved by backwards induction. As of the start of the third period, each worker's work history (i.e. sequence of prior employers) is publicly observed by all firms. Since employers are identical in this model, the work history can be thought of simply as revealing whether or not the worker's first and second-period employers differ.[34]

First, suppose that the worker switched jobs after the first period. Then a new match quality, η_2, is revealed (along with the worker's ability, θ) to the new matched pair at the end of the second period, whereas competing firms observe neither ability nor match quality. Alternatively, suppose the worker stays at the end of the first period. Then both $\eta_1$and θ are already known to the matched pair as of the end of the second period, because they were learned at the end of the first period. At the start of the third period, each firm simultaneously posts three wage offers. A firm matched with a stayer in the second period posts a wage offer of $W_3(y_1, s)$ to its stayer, $W_3^o(s)$ to other stayers in the market, and $W_3^o(q)$ to quitters in the market. A firm matched with a quitter in the second period posts a wage offer of $W_3(y_2, q)$ to its quitter, and two other offers—$W_3^o(s)$ to stayers and $W_3^o(q)$ to other quitters in the market.

After third-period wage offers are posted, each worker chooses the firm that makes the highest offer and remains with the current employer in the event of a tie. If a stayer decides to stay with the incumbent firm again (becoming a "superstayer"), the third-period productivity is the same as the second-period productivity, i.e. $y_3 = y_2 = \theta + \eta_1$. Similarly, if a quitter stays (becoming a "quitter-stayer") then third-period productivity is $y_3 = y_2 = \theta + \eta_2$. If a worker joins a new employer in the third period, a new match quality, η_3, is drawn and is revealed to the new matched pair at the

[33] We use a different notation to distinguish this from the first-period wage offer under symmetric learning, which is V_1.

[34] That is, a move from Firm A to Firm B is interpreted by the market the same as a move from Firm B to Firm A.

end of the third period. The third-period productivity for the new match is $y_3 = \theta + \eta_3$.

To ensure the existence of an equilibrium, we impose two (sufficient but not necessary) further assumptions on the distribution of ability θ and match quality η, namely, $\bar{\theta} - \theta_L \leq \bar{\eta} - \eta_L$ and $\bar{\theta} + \bar{\eta} \leq (\theta_H + \theta_L + \eta_H + \eta_L)/2$. The first inequality basically says that the range of the ability distribution should be no larger than that of the match-specific productivity distribution, which is reasonable as the latter could be unbounded. The second inequality requires that the distribution of the summation of ability and match quality cannot be severely negatively skewed. Examples of distributions that satisfy these two assumptions are the uniform, left-truncated normal, and exponential distributions with $\bar{\theta} \leq \bar{\eta}$, among others.

At the end of the first period, a worker with first-period output y_1 quits the incumbent firm if and only if

$$W_2(y_1) + \max\{W_3(y_1, s), W_3^o(s)\} < W_2^o + E_{\eta_2}[\max\{W_3(y_2, q), W_3^o(q)\}] \; [IR], \tag{9}$$

where the left-hand side is the sum of wages from the second and third periods if the worker stays, while the right-hand side is the expected sum of wages if the worker quits, and the expectation is with respect to the distribution of η_2. We denote this condition as IR and its complement as $\overline{IR}$.[35]

At the end of the second period, for stayers there exists a threshold value, $\hat{y}_2(s)$, satisfying

$$W_3^o(s) = \hat{y}_2(s) \Longrightarrow \hat{y}_2(s) = E\left[\theta | y_1 < \hat{y}_2(s), \overline{IR}\right] + \bar{\eta}, \tag{10}$$

such that the stayer quits if and only if $y_1 < \hat{y}_2(s)$. The incumbent firm offers

$$W_3(y_1, s) = \min\{y_1, \; \hat{y}_2(s)\}, \tag{11}$$

and the accepted incumbent wage offer, $\tilde{W}_3(y_1, s)$, is the same as the market offer, i.e. $\tilde{W}_3(y_1, s) = W_3^o(s) = \hat{y}_2(s)$.

Similarly, for quitters there exists a threshold value, $\hat{y}_2(q)$, satisfying

$$W_3^o(q) = \hat{y}_2(q) \Longrightarrow \hat{y}_2(q) = E[\theta | y_2 < \hat{y}_2(q), IR] + \bar{\eta}, \tag{12}$$

such that the quitter quits if and only if $y_2 < \hat{y}_2(q)$. The incumbent firm offers

$$W_3(y_2, q) = \min\{y_2, \; \hat{y}_2(q)\}, \tag{13}$$

and the accepted incumbent wage offer, $\tilde{W}_3(y_2, q)$, is same as the market offer, i.e. $\tilde{W}_3(y_2, q) = W_3^o(q) = \hat{y}_2(q)$.

The IR condition becomes

$$W_2(y_1) + \hat{y}_2(s) < W_2^o + \hat{y}_2(q) \; [IR]. \tag{14}$$

[35] Replace $<$ with $\geq$ in IR to get the complement.

At the end of the first period, the incumbent firm offers the wage $W_2(y_1) \leq y_1 + (y_1 - \hat{y}_2(s)) \cdot \mathbf{1}\{y_1 \geq \hat{y}_2(s)\}$. Therefore, there exists a threshold value, $\hat{y}_1$, satisfying

$$\hat{y}_1 + (\hat{y}_1 - \hat{y}_2(s)) \cdot \mathbf{1}\{\hat{y}_1 \geq \hat{y}_2(s)\} = W_2^o + \hat{y}_2(q) - \hat{y}_2(s), \tag{15}$$

such that the worker quits if and only if $y_1 < \hat{y}_1$ at the end of the first period. The incumbent firm offers

$$\begin{aligned} W_2(y_1) &= \min\{y_1 + (y_1 - \hat{y}_2(s)) \cdot \mathbf{1}\{y_1 \geq \hat{y}_2(s)\}, \\ &\qquad \hat{y}_1 + (\hat{y}_1 - \hat{y}_2(s)) \cdot \mathbf{1}\{\hat{y}_1 \geq \hat{y}_2(s)\}\}, \end{aligned} \tag{16}$$

and the accepted incumbent wage offer is

$$\tilde{W}_2(y_1) = \hat{y}_1 + (\hat{y}_1 - \hat{y}_2(s)) \cdot \mathbf{1}\{\hat{y}_1 \geq \hat{y}_2(s)\}. \tag{17}$$

The zero-profit condition at the end of the first period determines the wage W_2^o, i.e.

$$W_2^o = E[\theta | y_1 < \hat{y}_1] + \bar{\eta} + E\{[y_2 - \hat{y}_2(q)] \cdot \mathbf{1}(y_2 \geq \hat{y}_2(q)) | y_1 < \hat{y}_1\}. \tag{18}$$

Equations (10) and (12) become

$$\hat{y}_2(s) = E[\theta | \hat{y}_1 \leq y_1 < \hat{y}_2(s)] + \bar{\eta}, \tag{19}$$

$$\hat{y}_2(q) = E[\theta | y_2 < \hat{y}_2(q), y_1 < \hat{y}_1] + \bar{\eta}. \tag{20}$$

Equations (15), (19), and (20) characterize an equilibrium, which is summarized in the following proposition.

Proposition 2: *There exists at least one group of threshold values,* $(\hat{y}_1, \hat{y}_2(s), \hat{y}_2(q)) \in [\theta_L + \eta_L, \theta_H + \eta_H] \times [y^*, \bar{\theta} + \bar{\eta}] \times [\theta_L + \eta_L, y^*]$*, satisfying equations (15), (19), (20), and* $\hat{y}_1 < \hat{y}_2(s)$*, where* y^* *is the unique solution to*

$$y^* = E[\theta | y_1 < y^*] + \bar{\eta}. \tag{21}$$

At the end of the first period, a worker stays if and only if $y_1 \geq \hat{y}_1$*; at the end of the second period, a stayer stays again if and only if* $y_1 \geq \hat{y}_2(s)$*, and a quitter stays if and only if* $y_2 \geq \hat{y}_2(q)$*.*

At the end of the first period, the market offer is characterized by equation (18), the incumbent offer by equation (16), and the accepted incumbent offer by equation (17). At the end of the second period, the market offers are $W_3^o(s) = \hat{y}_2(s)$ *for stayers and* $W_3^o(q) = \hat{y}_2(q)$ *for quitters. The incumbent offers are characterized by equation (11) for stayers and equation (13) for quitters. The accepted offers match the market offers.*

The first-period wage is determined by the zero-profit condition,

$$W_1 = \bar{\theta} + \bar{\eta} + E\{[(y_1 - \hat{y}_1) + (y_1 - \hat{y}_2(s)) \cdot \mathbf{1}(y_1 \geq \hat{y}_2(s))] \cdot \mathbf{1}(y_1 \geq \hat{y}_1)\}. \tag{22}$$

Competing firms observe *two* turnover signals from each worker. Focusing on stayers, the subgame involving only the second and third periods is identical to a two-period model where competing firms observe only one signal from each worker at the end of the second period, namely the turnover decision (stay or quit).[36] That is, each stayer sends one signal to competing firms, and competing firms observe only one signal from each stayer. During the third period, stayers who quit are paid their expected productivity such that the firm's expected profit is zero. Incumbent firms match the outside option up to the worker's productivity level, so that stayers are paid the same during the third period regardless of whether they stay or quit.[37] The same reasoning applies to quitters. The difference between stayers and quitters is the distribution of their abilities, conditional on their earlier decisions to stay or quit. Stayers are expected to have higher abilities, on average, than quitters. Given that quitting sends competing firms a signal that the worker has a higher probability of being a lemon, workers with relatively high productivity (i.e. $\hat{y}_1 \leq y_1 < \hat{y}_2(s)$) will choose to stay at the end of the first period and quit at the end of the second period. They do this to differentiate themselves from those with lower productivity (i.e. $y_1 < \hat{y}_1$) who would choose to quit at the end of the first period. Thus, workers use staying (at the end of the first period) as a signaling device. This allows firms to sort workers into different groups (stayers or quitters). Therefore, the firm's belief is consistent, as required in equilibrium.

Furthermore, the first turnover decision cannot be perfectly inferred from the current labor market experience and current tenure. For example, those who quit at the end of the second period and work for new firms during the third period all have the same current experience (three periods) and current tenure (one period), but they potentially differ in their work histories (i.e. some may be stayers and others quitters). Thus, the number of prior employers plays an important independent role in determining the worker's current wage during the third period.

3.3 Testable Implications

We now turn to the empirical content of the preceding analyses. The following proposition states the main result summarizing the signaling effect of work history on the current wage.

Proposition 3: *Under symmetric learning, third-period wages are independent of work history. Under asymmetric learning, stayers are paid more than quitters even when they have identical labor market experience and current job tenure, i.e.* $W_3^o(s) > W_3^o(q)$.

The following example illustrates the proposition. Consider two workers, A and B, with the same ability, but suppose worker A's first-period match quality is higher than worker B's. Worker A's match quality is such that he is indifferent between staying and quitting; given our tie-breaking rule, he stays. This implies worker B quits. So by the end of the second period, worker A was continuously employed by one firm, whereas worker B has worked for two different firms. Suppose that both workers' second-period match qualities are such that they quit at the end of the second

[36] This situation corresponds to the upper branch of figure 2. Such a two-period model is a standard "lemons" model.

[37] Although we omit them for tractability, if we include random utility shocks as in Acemoglu and Pischke (1998) and Schönberg (2007), incumbent offers will in general vary with productivity. Including such shocks would not affect our main results.

period to join new third-period employers. Even though both workers have the same ability, the same observed current job tenure (i.e. one period with their respective third-period employers), and the same labor market experience (i.e. three periods), the market views these two workers differently. Conditional on the observed job history, worker B (the job hopper) is expected to have lower ability than worker A, so worker A is offered a higher third-period wage than worker B.

The intuition for this result is that prospective employers, when deciding what wage to offer a given worker, cannot discern to what extent the worker's history of job hopping is due to low match quality (which would not justify a lower wage offer) versus low ability (which *would* justify a lower wage offer). If a prospective employer knew for sure that a worker job hopped because of low match quality, then since match qualities are independent across firms, the employer would not downgrade the worker's wage offer. Similarly, if a prospective employer knew that a worker job hopped because of low ability, the employer would clearly downgrade the wage offer. But since the prospective employer cannot observe which of these two cases applies, the natural inclination is to assign some weight to the possibility that the job hopping occurred because of low ability, so the result is a wage penalty for job hopping.

The model's first and main testable implication follows from proposition 3 and is based on a difference between the two types of learning in the marginal effect of job hopping on the third-period wage, controlling for labor market experience and current job tenure. If learning is asymmetric, that marginal effect is negative, implying a wage penalty for job hopping. If learning is symmetric, that marginal effect is not identified if the model is taken literally. That is, due to $\eta^*_{s1} > \bar{\eta}$, there is no situation in which one stays at the end of the first period but quits at the end of the second period. Therefore, current job tenure and number of prior employers are collinear and their effects cannot be simultaneously identified given labor market experience.[38] Strictly speaking, this means that the marginal effects of job hopping under asymmetric and symmetric learning cannot be compared.

However, a reasonable generalization of the model resolves the identification problem while leaving our results qualitatively unchanged, and we assume it henceforth. Suppose that with a small positive probability the worker-firm match dissolves regardless of productivity (perhaps because a worker's spouse finds a good job in another location), as in Gibbons and Katz (1991), Greenwald (1986), and Waldman (2013), or the worker gets a new draw of match quality even without switching jobs (e.g. the worker might get a new boss or be assigned to a new task), as in Acemoglu and Pischke (1998) and Schönberg (2007). Then in this case, in the third period, a current tenure of one could map to either one or two previous firms, and the marginal effect of job hopping is negative under asymmetric learning and zero under symmetric learning, as discussed previously. Our results in table 2 are consistent with asymmetric learning for college graduates

[38]When learning is symmetric, in the third period, current tenure of one maps to two previous firms; current tenure of two maps to one previous firm; current tenure of three maps to zero previous firms. In the asymmetric learning model, however, current tenure of one in the third period could map to one (for stayers) or two (for quitters) previous firms.

and symmetric learning for high school graduates.[39]

A consequence is that if a current log-wage regression includes current job tenure, labor market experience, and job hopping measures on the right-hand side, dropping the job hopping variables should affect the coefficients on the other two variables under asymmetric learning (due to omitted variable bias) but not under symmetric learning. It should be noted that measures of labor market experience and current job tenure are regularly included in wage regressions, with researchers' interest focusing on the returns to both. Such regressions typically do not include prior job hopping on the right-hand side. Our models predict that such omissions are not a problem under symmetric learning but lead to bias in the coefficients of interest when learning is asymmetric. Results in table 5 show that omitting work history from the right-hand side leads to underestimates of the effects of experience and overestimates of the effects of current tenure, and these differences are statistically significant for college graduates but not for high school graduates. This, again, is consistent with our main finding that learning is asymmetric for college graduates but symmetric for high school graduates.

As one's labor market experience progresses, the work history contains more and more information, and job hopping measured by the number of previous jobs is just one piece of information. Alternatively, job hopping may be measured by the length of job tenures with previous firms. Returning to the two-worker example, recall that worker A (the stayer who has previous tenure of two periods) gets a higher third-period wage offer than the job-hopping worker B (the quitter who has previous tenures of one period each). This results in a positive relationship between the lengths of previous job tenures and one's current wage, after controlling for current tenure and labor market experience. In our robustness checks of table 3 that are based on an alternative measure of work history, we find a positive correlation between the current wage and the length of previous tenures, measured by the shortest prior tenure, for college graduates but not for high school graduates. We also find there exists a critical tenure cutoff (of one year) for college graduates. Previous tenures shorter than one year induce a larger penalty (figure 1). This is consistent with asymmetric learning for college graduates.

The second testable implication concerns employers' tendency to be sympathetic when there is reason to believe a job hopper is likely to have separated because of a bad match (as opposed to low ability). For example, even though prospective employers cannot observe whether job hopping occurred because of low ability or because of a bad match, the relative likelihood of the latter increases if the match was formed during an economic downturn. The wage penalty attached to job hopping under asymmetric learning should therefore be lower when workers enter the labor market during a bad year. This will not be true under symmetric learning, because in that case job hopping is predicted to have no effect on the current wage anyway (regardless of macroeconomic conditions at the point of labor market entry). Consistent with asymmetric learning, in table 4 we found a positive and statistically significant estimate of ξ for rookie college

[39]The finding of a smaller negative effect of job hopping for veteran college graduates could arise because less is learned from a new observation when there are already many prior observations.

graduates, revealing that the wage penalty attached to job hopping for this group is mitigated when workers enter the labor market during a bad economy. The corresponding effect was very small for high school graduates.

The preceding testable implication represents one application of a more general result from the theoretical model. Under asymmetric learning, the extent to which the market punishes job hopping should hinge on the relative dispersions of the match quality and ability distributions, both of which are publicly observed. Consider the extreme cases. When the dispersion of match quality is very small, employers know that an instance of job hopping was almost certainly due to low ability, and a large wage penalty results. Alternatively, when the dispersion of ability is very small, employers know that an instance of job hopping was almost certainly due to a bad match. Our evidence is consistent with the variance of match quality being high (low) relative to the variance of ability for high school (college) graduates.

Our model under asymmetric learning offers an explanation for the persistent negative effect of graduating from college in a recession. We find that when the unemployment rate is high, job tenure is generally shorter and is more likely to be the shortest job tenure in one's work history.[40] However, even if employers recognize this business cycle effect on job hopping and are sympathetic to it, they cannot perfectly separate the aggregate shock from other factors specific to the worker. Thus, if a worker's employment history includes a bad start that occurred during a recession, it would have a persistent effect on the worker's wage through job hopping stigma, even though match qualities dissolve upon separation and are independent across employers.

Although they cannot be addressed in the NLSY, there are additional implications related to recruitment and screening methods that employers use to obtain information about ability before hiring workers. If such methods were foolproof, they would mimic the symmetric learning world, i.e. a worker's job history would be irrelevant if the pre-hire screening methods could reveal ability perfectly. Although the methods are not foolproof, they presumably bring employers who use them one step closer to the symmetric learning world. Thus, a testable implication is that the interaction between the use of such methods and the number of prior employers should have a positive coefficient (i.e. mitigating the wage penalty from job hopping).

4 Discussion

Our results for college graduates are consistent with prior work that has found empirical support for asymmetric learning for that group (Schönberg, 2007; Pinkston, 2009; Kahn, 2013; Kim

[40]When we regress the length of each terminated tenure on the unemployment rate of the year when this tenure was terminated, the coefficient is -0.06 and is statistically significant at the 1% level, controlling for individual fixed effects, experience and its square, and age and its square. That is, when one job terminates in one year, the length of its tenure is shorter if the macroeconomic condition of the labor market in that year is worse. We also construct a dummy with 1 denoting that one job has the shortest tenure during one worker's entire work history in our sample, and regress this dummy variable on the unemployment rate, controlling for individual fixed effects. The coefficient is 0.036 and is significant at the 1% level. This implies that when a job ends in one year, the probability of it being the shortest tenure increases when labor market conditions worsen.

and Usui, 2014). Specifically, our findings are consistent with Schönberg (2007) and Kim and Usui (2014), both of which find that learning is asymmetric for college graduates but symmetric for workers with high school degrees only. However, as discussed in DeVaro and Waldman (2012), most jobs are probably better described by a blend of asymmetric and symmetric learning than by either perspective by itself. The job hopping stigma for college graduates can be explained by asymmetric learning. On the other hand, as in Altonji and Pierret (2001) and Arcidiacono et al. (2010), we also find that for high school graduates the positive marginal effect of the AFQT score increases with experience (tables 2 and 3), consistent with symmetric learning. Therefore, a reasonable interpretation of our differential findings between college and high school graduates is that the asymmetric learning mechanism is relatively stronger for college graduates, so that predictions of our theoretical model (under pure asymmetric learning) are clear in the data, whereas for high school graduates asymmetric learning is a less powerful force and is virtually indetectable in the data.[41]

Our conclusion that asymmetric learning is a particularly powerful force for college graduates is consistent with prior literature, but an unanswered question is *why* employer learning differs in this way between the two worker groups. One interpretation is that workers with only a high school degree are more likely to stay in their local community, and the dense network of social connections developed over a long tenure in the same location approximates a symmetric learning environment. In contrast, college graduates are relatively more likely to enter the labor market in a geographic location different from where they grew up (and move around more throughout their careers), and the lack of a dense local social network might make the asymmetric learning perspective more relevant.

A related point is suggested by the implication of our theoretical model that the magnitude of the wage penalty is driven by the relative variances of match quality and ability. A greater variance in the ability distribution relative to the match quality distribution among college graduates than among high school graduates would explain the result. Intuitively, learning about each worker's ability from their work history is quicker if the *ex ante* distribution of ability has less dispersion. In the extreme, if the ability distribution is degenerate then it takes no time to learn about ability, which is equivalent to the symmetric learning model where there is no wage penalty for job hopping. This implies that it would take longer for the market to learn a college graduate's ability than a high school graduate's ability, if the former has higher variance in the ability distribution relative to the match quality distribution.

An implication of our analysis is that match quality, despite being independent across employers, has persistent effects on careers in an asymmetric learning setting even when a worker switches firms. Bad luck with a particular firm (via a poor match quality) sticks with a worker even after leaving that firm, because prospective employers cannot distinguish bad luck from low ability. In contrast, in our symmetric learning environment all employers in the market have the

[41]Even for high school graduates, however, the results are suggestive of asymmetric learning when macroeconomic conditions at the point of labor market entry are added to the regression.

same information about ability at all times, so separations are seen as arising only from bad luck and not from low ability, and previous match quality has no effect on current or future wages.

These arguments have welfare implications. If learning is symmetric, poor matches are temporary because workers have no incentive to preserve them. A poorly matched worker seeks a new job, and the match is replaced by a new draw from the distribution, which is independent of the worker's ability. But when learning is asymmetric, workers who switch firms are adversely selected, and turnover is inefficient. At the end of the first period, workers quit if and only if $y_1 < \hat{y}_1$. High-ability workers with poor match qualities choose to stay so as to differentiate themselves from low-productivity lemons; and low-ability workers with high match qualities choose to quit as long as competing firms pay more. Both scenarios result in inefficiency and welfare loss.[42]

A further result is that the welfare loss increases with the number of periods, based on the following logic. From an efficiency standpoint, workers with poor matches in the first period should separate. But doing so sends a negative signal to the market, which has negative consequences for future earnings. The longer the future horizon, the greater the magnitude of those losses in future earnings, so the more tempted a poorly-matched worker will be to stay with the original employer in an effort to signal high ability to the market.[43] An implication is that the inefficiency should be greatest for early-career workers. A late-career worker who falls into a poor match has less to lose from ending it than would a new entrant with an entire career ahead of her.

5 Alternative Explanations

We now evaluate five potential alternative explanations for our empirical results that a) there is a job hopping wage penalty for college graduates but not high school graduates, and b) the wage penalty for college graduates is mitigated when workers enter the labor market during economic downturns.

5.1 On-the-job Search

In a standard on-the-job search model as in Burdett (1978) and Burdett and Mortensen (1998), workers search for and switch jobs in pursuit of higher wages, moving only when such offers appear. In general, such models predict a *zero* correlation between the current log-wage and the number of previous jobs, after controlling for labor market experience. Specifically, in a standard on-the-job search model with observable worker ability,[44] the wage in the current period t is the

[42]Not all asymmetric employer learning implies welfare loss. In Greenwald (1986), where the secondary market is sustained by an exogenous separation of worker-firm matches, turnover does not have any welfare consequences. Asymmetric employer learning in this context only shifts the utility allocation between workers and firms. The model in Pinkston (2009) has similar implications. By assuming that turnover creates a loss of productivity or some transaction costs, the model in Gibbons and Katz (1991) implies that turnover always causes welfare loss. The idiosyncratic utility shock in Schönberg (2007) also implies some welfare loss caused by asymmetric employer learning.

[43]Thus, the inefficiency in the period-1 turnover rate under asymmetric learning (i.e. the extent to which it falls short of the efficient period-1 turnover rate in the benchmark case of symmetric learning) is larger than the corresponding inefficiency in the period-2 turnover rate, and so on.

[44]The case in which ability is unobservable is described by our main model.

maximum wage offer received so far, i.e. $w_t = \theta + \max \{\eta_j, j = t_0, ..., t\}$, where t_0 is the first period since labor market entry (Burdett, 1978) or since the end of the last unemployment spell (Burdett and Mortensen, 1998). The match quality in period j, η_j, is a random draw from the common distribution, $G(\eta)$. Therefore, the wage w_t is a random variable with the cumulative distribution function (CDF) of $G(w_t - \theta)^{t-t_0+1}$. After controlling for ability θ, and labor market experience or the continuous employment spell $t - t_0$, the conditional mean of the wage w_t is independent of all other measures including the number of previous jobs.

In a variant of the search model with heterogeneous firm productivity as in Postel-Vinay and Robin (2002), the wage is a function of the worker's (observable) ability, the incumbent firm's productivity, and the distribution of productivity of all firms in the market. It is independent of the number of previous jobs. Thus, an alternative model based on on-the-job search is inconsistent with our main result that college graduates suffer a wage penalty from job hopping.

5.2 Public Ability with High-Ability Workers Infrequently Changing Jobs

Suppose that ability is publicly observed but that higher-ability workers are less likely to change jobs. Then job changers are, on average, lower-ability (and hence get paid less) even though ability is perfectly observed. This argument, which was suggested to us as a potential alternative explanation for the result that college graduates (but not high school graduates) suffer a wage penalty from job hopping, is unappealing for at least three reasons. First, it assumes that (for whatever reason) high-ability workers are less likely to be job changers, whereas in our model this is a result rather than an assumption. Second, the argument fails when ability is held constant. Controlling for ability, under the alternative explanation current wages should not depend on prior work history, conditional on work experience and current job tenure. Recall that AFQT is a control variable in our regressions. One might argue that AFQT is an imperfect, error-laden measure of ability, so that the alternative explanation is still somewhat viable. But in that case, when AFQT is dropped as a control we should observe a strengthening of our main result, which in fact does not happen, as seen in table 6. Third, in contrast to our asymmetric learning argument, the alternative does not offer an obvious explanation for why the job-hopping wage penalty for college graduates is mitigated when workers enter the labor market during economic downturns.

5.3 Symmetric Learning About Workers' Tastes for Staying

Prior work history might serve as a signal even in a symmetric learning model. Suppose that employers (including, perhaps, the incumbent employer) do not know how likely a worker is to invest in a long-term employment relationship until they observe the worker's behavior. The current employer might learn about that characteristic (i.e. the worker's "tastes for staying") at the same time as the rest of the market, by observing the worker's decisions to quit or stay. Thus, prior work history is used as a signal even though there is no asymmetric employer learning. Prior job tenures are correlated with current wages because long tenures are seen by all employers as suggesting stability, so they are all willing to pay extra for that.

This argument assumes there is an individual characteristic—taste for staying or taste for job hopping—in the error term of regression model (1) that affects both work history and current wages. Employers might not know the characteristic immediately but can learn about it, or might observe it from the beginning. In either case, if econometricians do not observe it there is an endogeneity problem. Finding an instrumental variable for work history is challenging. Macroeconomic conditions (either contemporaneously or at the point of labor market entry) offer one potential candidate. However, prior work (and indeed our own table 4) has found that macroeconomic conditions such as unemployment rates have direct effects on wages through various channels—human capital accumulation, search (Kahn, 2010; Oreopoulos et al., 2012), or implicit contracts (Beaudry and DiNardo, 1991).

There are two reasons why a symmetric learning explanation for our results for college graduates is unappealing. First, if the "tastes for staying" argument applies to college graduates, it is unclear why it would not also apply to high school graduates. That is, even if the symmetric learning argument is accepted as the explanation for our results for college graduates, it leaves the results for high school graduates unexplained. Second, the asymmetric learning model is able to explain not only the main result from table 2 but also the fact that this result (for rookie college graduates) is mitigated when workers enter the labor market during an economic downturn. The symmetric learning argument does not offer an explanation for the latter result. The reason is that, under the symmetric learning argument, employers would need to believe that workers' "tastes for staying" are stronger when they happen to enter the labor market during an economic downturn. But that is implausible; it is expected that workers' "tastes for staying" should be unrelated to what the economy happens to look like when they finish college and enter the labor market.

5.4 Fixed Costs of Hiring

One of our critiques of the preceding alternative explanation involving symmetric learning is that it is unclear why the argument would not hold for high school graduates as well as college graduates. The introduction of fixed costs of hiring (that vary by education level) responds to that critique. To elaborate, suppose learning about ability is identical for both college graduates and high school graduates but that there is a third random variable, "taste for variety," i.e. workers who value variety get bored with their current jobs after awhile and want to move on. This can be thought of as a "taste for job hopping," independent of ability and match quality, as in the preceding subsection. But now suppose that fixed costs of hiring (e.g. recruitment, screening, training, etc.) are higher for high-skilled workers than for low-skilled workers, so that turnover costs are higher for high-skilled workers. Employers may then interpret prior job hopping as a signal of short expected future tenure and high expected turnover costs. Given the high fixed costs of hiring a high-skilled worker, the employer may self-insure against the high likelihood of future separation by lowering the current wage. This could potentially explain our main result.

The preceding argument can be empirically distinguished from our employer learning explanation in the following way. From the employer's standpoint, high ability and low "taste for variety" both enhance productivity but for different reasons. High ability increases per-period productivity, whereas low "taste for variety" increases the expected number of periods of the employment relationship. An implication of "taste for variety" is that wage profiles should be steeper for job hoppers than for those who do not hop jobs, and this result should be particularly pronounced for college graduates, due to the high fixed costs of hiring them. The logic is that employers pass along more of the fixed hiring costs to job hoppers so as to insure against the possibility of future separation. Thus, such workers can expect to start initially at low wages, but once the fixed costs have been paid the wage will increase, so the likelihood of future separation is directly tied to the steepness of the wage profile.

In contrast, asymmetric learning about ability would not have this implication about the gradient of the wage profile. The reason is that a high-ability hire is more productive in *every period* after being hired. So there would be a wage penalty for job hoppers, but it would be persistent over the course of the new job tenure, because the job hopper would have lower expected ability in every period. Geometrically, compared to those who do not hop jobs, job hoppers suffer a parallel shifting down of their entire wage schedule (i.e. only the intercept changes). In contrast, if the wage penalty arises instead from a "taste for variety," this should involve a decrease in the intercept of the wage profile but an increase in its slope.

The preceding mechanism can be tested in the following regression model,

$$\ln w_{it} = X_{it}'\beta + WH_{it}'\gamma + \pi TEN_{it} + TEN_{it} \cdot WH_{it}'\xi + \epsilon_{it}, \tag{23}$$

where TEN_{it} represents current tenure; γ represents the intercept of the wage-tenure schedule, and ξ represents its slope, where both the intercept and slope can be influenced by job hopping. If the data generating process is governed by "taste for variety," then for job hoppers we should expect a negative γ and a positive ξ, given that WH_{it} is measured by the number of previous jobs.

Column 1 of table 7 displays estimation results for college graduates. The estimated γ is negative and statistically significant, but the estimated ξ (though positive) is far from statistically significant. The slope of the wage-tenure profile is insensitive to job hopping, which is inconsistent with the "taste for variety" mechanism. The results for high school graduates are displayed in column 2 and parallel those for college graduates, except that the wage penalty from job hopping is very small in magnitude (i.e. -0.003 versus -0.027 for college graduates). In short, the empirical evidence is inconsistent with the "taste for variety" mechanism. Furthermore, that mechanism's ability to explain our result that the job-hopping wage penalty for college graduates is mitigated when those workers enter the labor market during a bad economy is unclear.

5.5 Human Capital Accumulation

An alternative to a learning model is one in which workers accumulate firm-specific or general human capital, either by investing (Ben-Porath, 1967) or via learning-by-doing (Imai and Keane, 2004). Explanations for our main result that are based on human capital accumulation (either specific or general) are implausible given the controls in our empirical models. If human capital is entirely firm specific, each job change makes prior human capital irrelevant, and workers begin their next job with a clean slate. Therefore, a worker's human capital level (and wage) would be affected by current job tenure but not by prior job tenures.

An alternative that was suggested to us is that human capital is entirely general and that it is accumulated convexly during a job tenure. One interpretation of convex accumulation is that a worker must spend a sufficient amount of time with an employer to learn how the job is done, similar to a fixed cost. The current wage would then depend on the lengths of prior job tenures and, therefore, on the number of prior employers for a given amount of labor market experience. Only workers with sufficiently long tenures with their prior employers would have accumulated the productivity-enhancing skills that would make them more valuable to their current or potential employers. Further suppose that this convex accumulation occurs differently for college graduates versus high school graduates (in particular suppose the former group receives more training). Under these conditions, a worker's current wage may be negatively related to the number of prior employers, controlling for labor market experience. That is, convex general human capital accumulation can generate our first testable implication.

However, it is unclear that a convex accumulation function of entirely general human capital offers a realistic description of the typical case. Even setting that issue aside, it is unlikely that such accumulation could explain our result that the wage penalty for job hopping is mitigated when the worker enters the labor market in a bad economy. In the absence of the number of prior employers as an independent variable, the labor market conditions at entry would of course matter, because it would be a proxy for the average length of prior employment spells (the worse the labor market conditions, the shorter the spells and therefore the greater the number of prior employers). But once we control for the number of prior employers, macroeconomic conditions at the time of entry should be irrelevant.

6 Further Evidence From Employer Callbacks

Although we have focused on wages in this study, job hopping can potentially affect other career-relevant labor market outcomes. Even before an employment relationship begins, a job applicant's probability of receiving a callback from a prospective employer in response to a submitted job application may be reduced if the applicant's résumé exhibits a history of job hopping.

Bertrand and Mullainathan (2004) conduct an experiment by sending out fictitious résumés (i.e. job applications) in response to help-wanted ads posted in newspapers. The purpose of the experiment was to determine whether (pseudo) applicants whose names are commonly associated with non-whites are called back by prospective employers at a lower rate than those whose names are commonly associated with whites. Each résumé in the experiment includes the number of prior jobs one has held. In this section, we use the same data to investigate the correlation between the probability of receiving an employer callback and the number of prior jobs listed on the résumé.[45] The results suggest job hopping penalties for employer callbacks—and in particular the difference between college graduates and high school graduates—that are similar to those we have documented for current wages.

We estimate the following logit model:

$$Prob\left(callback_i = 1\right) = Logit\left(X_i'\beta + njobs_i \cdot \gamma\right), \tag{24}$$

where $callback_i$ is a dummy variable equaling one if résumé i receives an employer callback. The vector X_i includes résumé and job ad characteristics. The independent variable $njobs_i$ represents the number of prior jobs listed on the résumé and is our main focus.[46]

Table 8 reports the logit results for various groups of résumés. For the whole sample, the effect of the number of jobs listed on the résumé on the likelihood of an employer callback is negative but statistically insignificant, as shown in column 1. However, the effect is negative and statistically significant for those with six or fewer years of experience listed on the résumé (column 2), while it becomes positive but insignificant for those with more than six years of experience (column 3).[47] One additional job listed on the résumé is associated with a 1.3 percentage point decrease in the probability of an employer callback for those rookie applicants.[48]

We then estimate same logit models separately for each education subsample, namely college graduates (columns 4-6), those with some college (columns 7-9), and high school graduates or dropouts (column 10).[49] The results for college graduates almost replicate those for the whole sample in the first three columns. The results for those with some college are somewhat different—the number of jobs listed on the résumé has a negative effect on the likelihood of receiving an employer callback, regardless of the experience group, with a larger effect for rookie applicants. However, it is only statistically significant for the whole subsample. For the subsample of high school graduates or dropouts, the effect is positive but statistically insignificant.

In summary, for those résumés with at least some college education, the number of prior jobs listed on the résumé is negatively related to the likelihood of an employer callback; that is especially so within the first six years of experience, but there is no such negative effect for those with a high school degree (or less). This pattern of results, particularly the differing results by education level, suggests that the pattern of evidence we have documented for current wages may extend to other career-relevant labor market outcomes. Although the theoretical framework we propose is focused on wages rather than callbacks, an extension of the model that incorporates recruiting

[45] We obtained the data from the supplement of Bertrand and Mullainathan (2004), downloaded directly from the journal website.

[46] On the 4870 résumés the number of jobs ranges from 1 to 7, with median 4, mean 3.661 and standard deviation 1.219. The correlation between the number of jobs and being black is essentially zero (-0.0025).

[47] The results are similar if the threshold is five years. Both are insignificant if the threshold is seven years.

[48] The mean of the callback dummy is 0.08 for the whole sample.

[49] Due to the sample size, we include both high school graduates and dropouts in the last education subsample. There are only 287 observations in this subsample, so we use the entire subsample to estimate the statistical model.

and screening costs should produce predictions consistent with the callback results.[50]

7 Conclusion

We have identified evidence from the NLSY79 of a job-hopping wage penalty for college graduates. The result is statistically and economically significant, larger for new entrants than for veterans, particularly large when jobs are held for less than one year, and mitigated when workers enter the labor market during an economic downturn. Although this wage penalty for job-hopping college graduates is a new result that is interesting in its own right, we have also shown that it can be interpreted as a test of asymmetric (versus symmetric) employer learning about worker ability. Evidence that job hopping leads to a wage penalty for college graduates is consistent with asymmetric learning for that group, and the near absence of such a relationship for high school graduates is consistent with symmetric learning for that group. That conclusion echoes those of Schönberg (2007) and Kim and Usui (2014) based on different arguments, and even though we use the NLSY79 data (as they do) to reach that conclusion, we do so via an entirely different approach.

Our analysis introduces the idea that work history is a signal of ability, which workers consider when making decisions to stay with or leave an employer. Because employers cannot discern whether a separation arose from a bad match or from low ability, they assign some weight to the latter possibility, which negatively affects job hoppers' wages. Thus, even though match quality is independent across firms, workers carry the ill effects of a bad match with them even after separating. This causes poorly matched workers to think twice before ending a match. It also means early-career workers have the most to lose by prematurely ending a bad match, since they have an entire career ahead of them in which to suffer wage penalties imposed by the market.

The insights of this analysis have application beyond the labor market. For example, consider two recently divorced men re-entering the marriage market. Both have been married for nine years, but one was married to the same spouse for all nine years whereas the other was married to three different spouses, each for three years. For both men, the prior divorces may have been caused by bad luck (i.e. poor match quality) or because they were bad husbands. But since prospective spouses cannot distinguish between those cases, the man with three divorces in his past will be viewed more skeptically than the once-divorced man, in the eyes of prospective spouses. Since the thrice-divorced man has lower market value than the once-divorced man, the expected quality of his subsequent spouse is reduced. A further implication is that even if it is apparent to both parties to a new marriage that the match is horrible, it may be wise to hold off on a divorce "for a respectable amount of time" to mitigate the negative signal resulting from divorce, thereby improving the prospects of a better marriage in the future.[51]

[50] We leave it for future research.

[51] To continue the analogy, our model predicts that under asymmetric learning employers should be more forgiving when there is good reason to believe that job hoppers faced bad matches in their employment histories. Similarly, in the marriage market, prospective spouses should be more forgiving of divorcees who were previously in arranged (rather than voluntary) marriages.

There are some elements we have not addressed that, if incorporated into our theoretical model in future research, could generate further insights. One is workers' search behavior. Another is schooling choices, or other observable worker characteristics (other than work history) that relate to ability. In such an extension, workers could signal ability to prospective employers both through their schooling choices and through their career histories. Another extension would allow for involuntary as well as voluntary turnover. Given that our model has no involuntary turnover, we do not have unemployment spells as in Pinkston (2009). However, our empirical results are robust to specifications that include continuous employment spells. We also do not allow for promotions, as studied in Waldman (1984), Bernhardt (1995), Owan (2004), DeVaro and Waldman (2012) and Prasad and Tran (2013), which empirically are an important source of wage growth and internal (i.e. within-firm) job mobility. Although extending the theory to incorporate promotions would be interesting, empirical testing would be infeasible in the NLSY79 data set, which has no information on internal promotions. It would be possible, however, in a large-scale, linked, employer-employee Finnish dataset (Kauhanen and Napari, 2012).

References

Acemoglu, Daron and Jorn-Steffen Pischke, "Why Do Firms Train? Theory and Evidence," *The Quarterly Journal of Economics*, Feburary 1998, *113* (1), 79–119.

Akerlof, George A., "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, August 1970, *84* (3), 488–500.

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

_ , Lisa B. Kahn, and Jamin D. Speer, "Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success," Working Paper 20531, National Bureau of Economic Research September 2014.

Antonovics, Kate and Limor Golan, "Experimentation and job choice," *Journal of Labor Economics*, 2012, *30* (2), 333–366.

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

Ariga, Kenn, Fumio Ohtake, Masaru Sasaki, and Zheren Wu, "Wage growth through job hopping in China," 2012.

Bagnoli, Mark and Ted Bergstrom, "Log-concave probability and its applications," *Economic Theory*, 2005, *26*, 445–469.

Beaudry, Paul and John DiNardo, "The Effect of Implicit Contracts on the Movement of Wages Over the Business Cycle: Evidence from Micro Data," *Journal of Political Economy*, August 1991, *99* (4), 665–688.

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

Bernhardt, Dan, "Strategic Promotion and Compensation," *The Review of Economic Studies*, April 1995, *62* (2), 315–339.

Bertrand, Marianne and Sendhil Mullainathan, "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *The American Economic Review*, 2004, *94* (4), 991–1013.

Bills, David B, "Employers' Use of Job History Data for Making Hiring Decisions: A Fuller Specification of Job Assignment and Status Attainment," *Sociological Quarterly*, Spring 1990, *31* (1), 23–35.

— , "Labor market information and selection in a local restaurant industry: The tenuous balance between rewards, commitments, and costs," in "Sociological Forum," Vol. 14 Springer December 1999, pp. 583–607.

Bognanno, Michael and Eduardo Melero, "Promotion Signals, Experience and Education," *Journal of Economics and Management Strategy*, 2014. forthcoming.

Buchinsky, Moshe, Denis Fougere, Francis Kramarz, and Rusty Tchernis, "Interfirm mobility, wages and the returns to seniority and experience in the United States," *The Review of Economic Studies*, 2010, *77* (3), 972–1001.

Burdett, Kenneth, "A Theory of Employee Job Search and Quit Rates," *The American Economic Review*, 1978, *68* (1), pp. 212–220.

— **and Dale T. Mortensen**, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, May 1998, *39* (2), 257–273.

Cassidy, Hugh, Jed DeVaro, and Antti Kauhanen, "Promotion Signaling: New Theory and Evidence from Multiple Firms," 2015. Working paper.

DeVaro, Jed and Hodaka Morita, "Internal promotion and external recruitment: a theoretical and empirical analysis," *Journal of Labor Economics*, 2013, *31* (2), 227–269.

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

— **, Suman Ghosh, and Cindy Zoghi**, "Job characteristics and labor Market discrimination in promotions: New theory and empirical evidence," 2015. Working paper.

Fallick, Bruce, Charles A Fleischman, and James B Rebitzer, "Job-hopping in Silicon Valley: some evidence concerning the microfoundations of a high-technology cluster," *The Review of Economics and Statistics*, 2006, *88* (3), 472–481.

Farber, Henry S, "Employment insecurity: The decline in worker-firm attachment in the united states," 2008.

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

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

— **and Michael Waldman**, "A Theory of Wage and Promotion Dynamics inside Firms," *The Quarterly Journal of Economics*, 1999, *114* (4), pp. 1321–1358.

— **and** — , "Enriching a Theory of Wage and Promotion Dynamics inside Firms," *Journal of Labor Economics*, 2006, *24* (1), 59–108.

Golan, Limor, "Wage Signaling: A Dynamic Model of Intrafirm Bargaining and Asymmetric Learning," *International Economic Review*, August 2009, *50* (3), 831–854.

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

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

Heckman, James J. and Bo E. Honore, "The Empirical Content of the Roy Model," *Econometrica*, September 1990, *58* (5), 1121–1149.

_ , Jora Stixrud, and Sergio Urzua, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, July 2006, *24* (3), 411–481.

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

Imai, Susumu and Michael P. Keane, "Intertemporal Labor Supply and Human Capital Accumulation," *International Economic Review*, May 2004, *45* (2), 601–641.

Johnson, William R., "A Theory of Job Shopping," *The Quarterly Journal of Economics*, May 1978, *92* (2), pp. 261–278.

Jovanovic, Boyan, "Job Matching and the Theory of Turnover," *The Journal of Political Economy*, October 1979, *87* (5), 972–990.

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

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

Kauhanen, Antti and Sami Napari, "Career and wage dynamics: Evidence from linked employer-employee data," *Research in Labor Economics*, 2012, *36*, 35–76.

Kim, Seik and Emiko Usui, "Employer Learning, Job Changes, and Wage Dynamics," 2014.

Lange, Fabian, "The speed of employer learning," *Journal of Labor Economics*, 2007, *25* (1), 1–35.

Lazear, Edward P., "Entrepreneurship," *Journal of Labor Economics*, October 2005, *23* (4), 649–680.

Milgrom, Paul and Sharon Oster, "Job discrimination, market forces, and the invisibility hypothesis," *The Quarterly Journal of Economics*, 1987, pp. 453–476.

Munasinghe, Lalith and Karl Sigman, "A hobo syndrome? Mobility, wages, and job turnover," *Labour Economics*, 2004, *11* (2), 191–218.

Nagypal, Eva, "Learning by Doing vs. Learning About Match Quality: Can We Tell Them Apart?," *Review of Economic Studies*, March 2007, *74* (2), 537–566.

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

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

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz, "The Short- and Long-Term Career Effects of Graduating in a Recession," *American Economic Journal: Applied Economics*, January 2012, *4* (1), 1–29.

Owan, Hideo, "Promotion, turnover, earnings, and firm-sponsored training," *Journal of Labor Economics*, 2004, *22* (4), 955–978.

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

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, November 2002, *70* (6), 2295–2350.

Prasad, Suraj and Hien Tran, "Work practices, incentives for skills, and training," *Labour Economics*, 2013, *23*, 66–76.

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

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

Waldman, Michael, "Job Assignments, Signalling, and Efficiency," *The RAND Journal of Economics*, Summer 1984, *15* (2), 255–267.

_ , "Up-or-Out Contracts: A Signaling Perspective," *Journal of Labor Economics*, April 1990, *8* (2), 230–250.

_ , "Asymmetric Learning and the Wage/productivity Relationship," *Journal of Economic Behavior & Organization*, December 1996, *31* (3), 419–429.

_ , "Classic promotion tournaments versus market-based tournaments," *International Journal of Industrial Organization*, 2013, *31* (3), 198–210.

Zhang, Ye, "Employer Learning Under Asymmetric Information: The Role of Job Mobility," 2011. Working paper.

Table 1: Summary statistics of NLSY79 sample (1979-2008).

	College Graduates		**High School Graduates**	
	Mean	S.D.	Mean	S.D.
No. individuals	923		2949	
No. observations	18544		61058	
Real hourly wage (2004$)	24.84	(13.332)	14.11	(7.425)
Number of all jobs[a]	6.634	(4.468)	6.516	(5.048)
Number of previous jobs[a]	5.634	(4.468)	5.516	(5.048)
Experience[b]	10.496	(6.925)	8.639	(7.078)
Current tenure[b]	4.051	(4.517)	3.125	(4.251)
AFQT[a]	1.089	(0.771)	-0.413	(0.792)
White[a]	0.75	(0.433)	0.567	(0.496)
Hispanic[a]	0.092	(0.289)	0.172	(0.377)
Black[a]	0.158	(0.365)	0.261	(0.439)
Years of schooling[a]	17.012	(1.384)	11.081	(1.422)
Age	33.595	(7.247)	28.943	(7.943)
Urban area	0.85	(0.357)	0.76	(0.427)
USA residency at age 14	0.986	(0.118)	0.977	(0.149)
Unemp rate at entry[a]	3.264	(1.751)	10.030	(3.170)

[a] Variables reported at the individual level. "High School Graduates" are those with 12 or fewer years of schooling; "College Graduates" are those with 16 or more years of schooling.
[b] Tenure and experience are divided by 52 (i.e. measured in years).

Table 2: Effects of work history on current log-wage.

	1	2	3	4	5	6
	College graduates only			**High school graduates only**		
	All	**Rookies**	**Veterans**	**All**	**Rookies**	**Veterans**
No. of previous jobs	-0.025***	-0.048***	-0.025***	-0.002	0.002	-0.002
	(0.005)	(0.011)	(0.006)	(0.002)	(0.004)	(0.002)
AFQT score	0.077***	0.056**	0.091***	0.026***	0.019**	0.043***
	(0.019)	(0.023)	(0.028)	(0.007)	(0.008)	(0.013)
AFQT×Experience	0.000	0.010	-0.001	0.005***	0.007**	0.003***
	(0.002)	(0.006)	(0.002)	(0.001)	(0.003)	(0.001)
R-squared	0.39	0.32	0.30	0.34	0.26	0.30
Observations	11,661	3,767	7,894	56,178	21,038	35,140

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust Huber/White standard errors are in parentheses, clustered at the individual level.
"Rookies" refer to the first five years in the labor market and "Veterans" refer to periods after the first five years. All regressions include experience and its square, current tenure and its square, years of schooling, (for college graduates) the interactions of years of schooling with experience and experience squared, Hispanic, Black, urban residency, USA residency at age 14, a dummy variable for being the interview year, dummy variables for years, (for college graduates) the interactions of year dummies with years of schooling, dummy variables for one-digit industry and occupation codes, and (for high school graduates) a dummy for finishing high school.

Table 3: Effects of work history on current log-wage, measured by the length of shortest prior tenure.

	1	2	3	4	5	6
	College graduates only			**High school graduates only**		
	All	**Rookies**	**Veterans**	**All**	**Rookies**	**Veterans**
Shortest previous tenure	0.011***	0.047***	0.010**	0.002	0.004	0.002
	(0.004)	(0.015)	(0.004)	(0.003)	(0.006)	(0.003)
AFQT score	0.076***	0.059**	0.092***	0.026***	0.019**	0.043***
	(0.020)	(0.024)	(0.028)	(0.007)	(0.008)	(0.013)
AFQT×Experience	0.001	0.009	-0.000	0.005***	0.008**	0.003***
	(0.002)	(0.007)	(0.002)	(0.001)	(0.003)	(0.001)
R-squared	0.38	0.31	0.29	0.34	0.25	0.30
Observations	11,661	3,767	7,894	56,178	21,038	35,140

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust Huber/White standard errors are in parentheses, clustered at the individual level.
All regressions include experience and its square, current tenure and its square, years of schooling, (for college graduates) the interactions of years of schooling with experience and experience squared, Hispanic, Black, urban residency, USA residency at age 14, a dummy variable for being the interview year, dummy variables for years, (for college graduates) the interactions of year dummies with years of schooling, dummy variables for one-digit industry and occupation codes, and (for high school graduates) a dummy for finishing high school.

Table 4: Effects of work history interacting labor market entry conditions on current log-wage.

	1	2	3	4	5	6	7	8
	College graduates				**High school graduates**			
	Rookies		**Veterans**		**Rookies**		**Veterans**	
Number of previous jobs	-0.134***	-0.422**	-0.025	-0.191*	-0.040***	-0.051*	-0.004	-0.015
	(0.037)	(0.206)	(0.025)	(0.107)	(0.011)	(0.029)	(0.006)	(0.017)
Number of previous jobs	0.035**	0.261*	-0.000	0.127	0.004***	0.006	0.000	0.002
$\times$Unemp rate at LM entry	(0.014)	(0.155)	(0.009)	(0.081)	(0.001)	(0.005)	(0.001)	(0.003)
Number of previous jobs		-0.042		-0.023		-0.000		-0.000
$\times$Unemp rate at LM entry^2		(0.028)		(0.015)		(0.000)		(0.000)
AFQT score	0.050**	0.050**	0.093***	0.093***	0.032***	0.032***	0.045***	0.045***
	(0.024)	(0.024)	(0.028)	(0.028)	(0.010)	(0.010)	(0.016)	(0.016)
AFQT$\times$Experience	0.012*	0.012*	-0.001	-0.001	0.005	0.005	0.003***	0.003***
	(0.007)	(0.007)	(0.002)	(0.002)	(0.004)	(0.004)	(0.001)	(0.001)
Unemp rate at LM entry	-0.048*	-0.053	-0.029	-0.360	-0.018***	-0.043***	-0.002	-0.004
	(0.028)	(0.303)	(0.045)	(0.438)	(0.002)	(0.012)	(0.004)	(0.024)
Unemp rate at LM entry^2		0.001		0.060		0.001**		0.000
		(0.055)		(0.081)		(0.001)		(0.001)
R-squared	0.31	0.31	0.29	0.29	0.22	0.22	0.29	0.29
Observations	3,751	3,751	7,861	7,861	14,998	14,998	24,159	24,159

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors in parentheses are Huber/White, clustered at the individual level.
"Rookies" refer to the first five years in the labor market and "Veterans" refer to periods after the first five years. All regressions include experience and its square, current tenure and its square, years of schooling, (for college graduates) the interactions of years of schooling with experience and experience squared, Hispanic, Black, urban residency, USA residency at age 14, a dummy variable for being the interview year, dummy variables for one-digit industry and occupation codes, and (for high school graduates) a dummy for finishing high school.

Table 5: Comparison of returns to experience and tenure, with or without including work history.

	1	2	3	4	5	6	7	8	9	10	11	12
			College graduates						**High school graduates**			
	All		**Rookies**		**Veterans**		**All**		**Rookies**		**Veterans**	
Include history?	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Experience	0.087***	0.070***	0.211***	0.141***	0.064***	0.052***	0.062***	0.059***	0.087***	0.086***	0.049***	0.048***
	(0.007)	(0.007)	(0.027)	(0.022)	(0.011)	(0.011)	(0.003)	(0.003)	(0.010)	(0.008)	(0.004)	(0.004)
Experience^2	-0.002***	-0.002***	-0.017***	-0.012***	-0.001***	-0.001***	-0.001***	-0.001***	-0.004***	-0.004***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.004)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
Tenure	0.001	0.011***	-0.019	0.027***	0.001	0.011***	0.013***	0.016***	0.042***	0.042***	0.013***	0.014***
	(0.003)	(0.003)	(0.012)	(0.010)	(0.003)	(0.003)	(0.001)	(0.001)	(0.006)	(0.004)	(0.001)	(0.001)
AFQT	0.081***	0.091***	0.080***	0.082***	0.081***	0.093***	0.065***	0.065***	0.035***	0.035***	0.085***	0.084***
	(0.018)	(0.020)	(0.018)	(0.018)	(0.022)	(0.024)	(0.007)	(0.007)	(0.007)	(0.007)	(0.010)	(0.010)
R-squared	0.39	0.38	0.31	0.30	0.30	0.28	0.34	0.33	0.25	0.25	0.29	0.29
Observations	11,661	11,661	3,767	3,767	7,894	7,894	56,178	56,178	21,038	21,038	35,140	35,140

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust Huber/White standard errors are in parentheses, clustered at the individual level.
"Rookies" refer to the first five years in the labor market and "Veterans" refer to periods after the first five years. Odd columns (1, 3,...11) include the number of previous jobs, while even columns (2, 4, ...12) do not. All regressions include years of schooling, (for college graduates) the interactions of years of schooling with experience and experience squared, Hispanic, Black, urban residency, USA residency at age 14, a dummy variable for being the interview year, dummy variables for years, the interactions of year dummies with years of schooling, dummy variables for one-digit industry and occupation codes, and (for high school graduates) a dummy for finishing high school.

Table 6: Effects of work history on current log-wage without including AFQT.

	1	2	3	4	5	6
	College graduates			**High school graduates**		
	All	**Rookies**	**Veterans**	**All**	**Rookies**	**Veterans**
Number of previous jobs	-0.027***	-0.050***	-0.027***	-0.002	0.003	-0.002
	(0.006)	(0.011)	(0.006)	(0.002)	(0.004)	(0.002)
R-squared	0.38	0.31	0.29	0.33	0.25	0.29
Observations	11,661	3,767	7,894	56,178	21,038	35,140

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust Huber/White standard errors are in parentheses, clustered at the individual level.
"Rookies" refer to the first five years and "Veterans" refer to periods after the first five years. Regressions are identical to those in table 2 but dropping the AFQT score and the interaction of the AFQT score with experience.

Table 7: Slope of return to current tenure.

	1	2
	College graduates	**High school graduates**
Number of previous jobs	0.000	0.001
×Current tenure	(0.001)	(0.000)
Number of previous jobs	-0.027***	-0.003*
	(0.007)	(0.002)
AFQT score	0.077***	0.026***
	(0.019)	(0.007)
AFQT×Experience	0.000	0.005***
	(0.002)	(0.001)
Current tenure	0.012*	0.033***
	(0.007)	(0.003)
Current tenure square	-0.001**	-0.001***
	(0.000)	(0.000)
R-squared	0.39	0.34
Observations	11,661	56,178

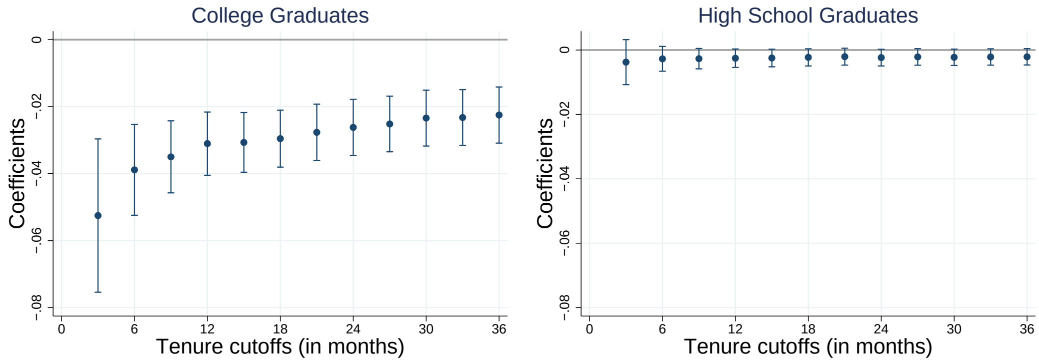
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust Huber/White standard errors are in parentheses, clustered at the individual level.
All regressions include experience and its square, years of schooling, the interactions of years of schooling with experience and experience squared, Hispanic, Black, urban residency, USA residency at age 14, a dummy variable for being the interview year, dummy variables for years, the interactions of year dummies with years of schooling, dummy variables for one-digit industry and occupation codes, and (for high school graduates) a dummy for finishing high school.

Table 8: Effect of the number of prior jobs on the likelihood of an employer callback[a]

	1	2	3	4	5	6	7	8	9	10
	The Whole Sample[a]			**College Graduates**			**Some College**			**High School**
	all	**exp<=6**[c]	**exp>6**[d]	**all**	**exp<=6**	**exp>6**	**all**	**exp<=6**	**exp>6**	**or less**[b]
Number of prior jobs	-0.007	-0.013*	0.003	-0.005	-0.014*	0.002	-0.025**	-0.014	-0.007	0.031
	(0.004)	(0.007)	(0.007)	(0.004)	(0.008)	(0.008)	(0.012)	(0.020)	(0.019)	(0.025)
Has employment gaps	0.022**	0.038**	-0.005	0.020	0.041*	-0.005	0.007	0.029	0.047	-0.077
	(0.011)	(0.017)	(0.018)	(0.013)	(0.023)	(0.020)	(0.026)	(0.060)	(0.035)	(0.077)
Black	-0.033***	-0.025***	-0.038***	-0.028***	-0.018*	-0.031**	-0.054***	-0.044**	-0.059**	-0.049
	(0.007)	(0.008)	(0.011)	(0.008)	(0.010)	(0.012)	(0.014)	(0.019)	(0.026)	(0.037)
Female	0.008	-0.017*	0.029	0.016	-0.008	0.037**	-0.067	-0.053	0.021	-0.089
	(0.011)	(0.010)	(0.018)	(0.010)	(0.011)	(0.018)	(0.052)	(0.070)	(0.062)	(0.054)
Experience	0.008***	0.037	0.019**	0.006*	0.011	0.016*	0.007	0.071	-0.017	0.033*
	(0.003)	(0.029)	(0.008)	(0.004)	(0.049)	(0.009)	(0.008)	(0.045)	(0.023)	(0.018)
Experience^2	-0.000**	-0.004	-0.001**	-0.000	-0.001	-0.000	-0.000	-0.010*	0.001	-0.001
	(0.000)	(0.003)	(0.000)	(0.000)	(0.005)	(0.000)	(0.000)	(0.006)	(0.001)	(0.001)
Observations	4,784	2,404	2,380	3,436	1,636	1,788	990	527	463	287

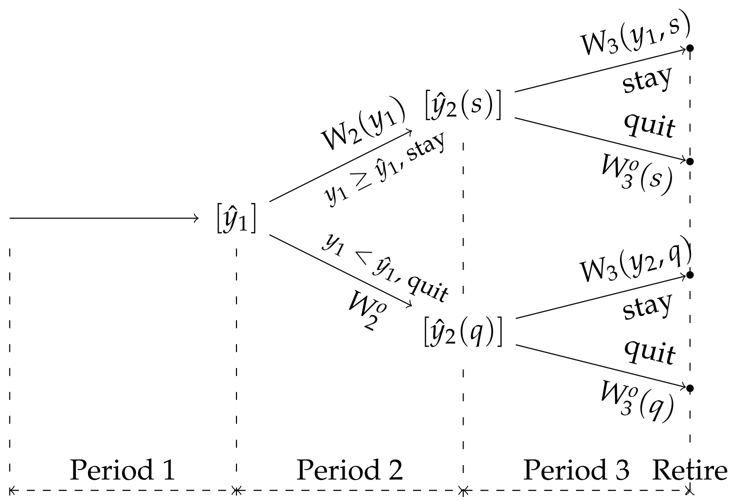
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors in parentheses are clustered at the employment-ad level.
Each column gives the results of a logit regression where the dependent variable is the callback dummy. Reported in the table are estimated marginal changes in probability for the continuous variables and estimated discrete changes for the dummy variables. All regressions include a city dummy, (for columns 1-3 only) education dummies or (for column 10 only) a dummy for graduating from high school, a vector of dummy variables for résumé's characteristics (volunteering, military, having an email address, working while at school, honors, computer skills, and special skills), a vector of dummy variables for job requirements as listed in the employment ad (experience, computer skills, communication skills, organization skills, and education), six occupation dummies, a dummy if the ad mentions it is an equal opportunity employer (EOE), and some demographic variables in the applicant's zip code (fraction of blacks, log median household income, fraction of high school dropouts, fraction of college graduates, and log per capita income).
[a]The entire set of sent résumés.
[b]The high school graduates and dropouts.
[c]Those with 6 or fewer years of experience.
[d]Those with more than 6 years of experience.

Figure 1: Effect (on the current log wage) of the number of prior jobs shorter than various tenure cutoffs.



Dots represent regression coefficients, and vertical lines represent 90% confidence intervals. For example, consider the second vertical line in the left panel (for college graduates), corresponding to a tenure cutoff of six months. In a log-wage regression, the coefficient of "number of previous jobs with tenure less than 6 months" is -0.04. The underlying regressions are the same as those in table 2, apart from the measures of work history (i.e. in the regressions underlying figure 1, the work history measure is the "number of previous jobs with tenure less than X months", where X is indicated on the horizontal axis.

Figure 2: The asymmetric learning model.



In the symmetric equilibrium, at the end of the first period the incumbent wage offer for a worker with productivity y_1 is $W_2(y_1)$, and the market wage offer is W_2^o. Workers with productivity less than $\hat{y}_1$ quit (and are called "quitters") and otherwise stay (and are called "stayers"). At the end of the second period, for a stayer with productivity y_1 the incumbent wage offer is $W_3(y_1, s)$, and the market wage offer is $W_3^o(s)$; for a quitter with productivity y_2 the incumbent wage offer is $W_3(y_2, q)$, and the market wage offer is $W_3^o(q)$. Among stayers, those with productivity less than $\hat{y}_2(s)$ quit and otherwise stay. Among quitters, those with productivity less than $\hat{y}_2(q)$ quit and otherwise stay. All workers retire at the end of the third period.

Appendix

A Proof of Proposition 1: Symmetric Learning Model

At the end of the second period, the zero-profit condition yields the market offer,

$$V_3^o(\theta) = \theta + \bar{\eta}.$$

The incumbent offer is

$$V_3(\theta, \eta_j) = \min\left\{\theta + \eta_j, V_3^o(\theta)\right\} = \theta + \min\left\{\eta_j, \bar{\eta}\right\}.$$

Therefore, a worker quits if and only if $\eta_j < \bar{\eta}$. The accepted incumbent offer, $\tilde{V}_3(\theta)$, is the same as the market offer, i.e. $\tilde{V}_3(\theta) = V_3^o(\theta)$.

At the end of the first period, the market offer, $V_2^o(\theta)$, is determined by the zero expected profit condition,

$$V_2^o(\theta) = \theta + \bar{\eta} + E\left\{(\eta_2 - \bar{\eta}) \cdot \mathbf{1}\left\{\eta_2 \geq \bar{\eta}\right\}\right\},$$

and the incumbent offer is

$$V_2(\theta, \eta_1) = \min\left\{\theta + \eta_1 + (\eta_1 - \bar{\eta}) \cdot \mathbf{1}\left\{\eta_1 \geq \bar{\eta}\right\}, V_2^o(\theta)\right\},$$

where the first argument in the min is the maximum wage that the incumbent employer is willing to offer. The worker will receive the same wage during the third period regardless of whether the worker stayed or quit at the end of the first period. Therefore, only the second-period wage affects turnover.

The preceding implies that a worker quits if and only if

$$\eta_1 + (\eta_1 - \bar{\eta}) \cdot \mathbf{1}\left\{\eta_1 \geq \bar{\eta}\right\} < \bar{\eta} + E\left\{(\eta_2 - \bar{\eta}) \cdot \mathbf{1}\left\{\eta_2 \geq \bar{\eta}\right\}\right\}.$$

Since the right-hand side is a constant and the left-hand side is a strictly increasing function of η_1, there exists a unique threshold value, η_{s1}^*, which solves

$$\eta_{s1}^* + (\eta_{s1}^* - \bar{\eta}) \cdot \mathbf{1}\left\{\eta_{s1}^* \geq \bar{\eta}\right\} = \bar{\eta} + E\left\{(\eta_2 - \bar{\eta}) \cdot \mathbf{1}\left\{\eta_2 \geq \bar{\eta}\right\}\right\}, \tag{A.1}$$

such that at the end of the first period a worker quits if and only if $\eta_1 < \eta_{s1}^*$. Note that $\eta_{s1}^* > \bar{\eta}$, since the second term on the right-hand side of equation (A.1) is positive.[52] Therefore, equation (A.1) becomes

$$\eta_{s1}^* = \bar{\eta} + \frac{1}{2} E\left\{(\eta_2 - \bar{\eta}) \cdot \mathbf{1}\left\{\eta_2 \geq \bar{\eta}\right\}\right\}.$$

[52] Intuitively, this is due to the convexity in the expected match quality, given the option of quitting again at the end of the second period.

Again, the accepted incumbent offer matches the market offer, i.e. $\tilde{V}_2(\theta) = V_2^o(\theta)$.

The first-period wage is determined by the zero expected profit condition given free entry to the labor market,

$$\begin{aligned} V_1 &= \bar{\theta} + \bar{\eta} + E\left\{[\eta_1 - \bar{\eta} - E\left\{(\eta_2 - \bar{\eta}) \cdot \mathbf{1}\{\eta_2 \geq \bar{\eta}\}\right\}] \cdot \mathbf{1}\{\eta_1 \geq \eta_{s1}^*\} \right. \\ &\quad \left. + (\eta_1 - \bar{\eta}) \cdot \mathbf{1}\{\eta_1 \geq \bar{\eta}\}\right\}. \blacksquare \end{aligned}$$

B Proof of Proposition 2: Asymmetric Learning Model

The proof has two steps. First we prove the existence and uniqueness of y^*, and second we prove the existence of $(\hat{y}_1, \hat{y}_2(s), \hat{y}_2(q))$.

In the first step, we apply the intermediate value theorem to prove the existence of y^*.

Define the right-hand side of Equation (21) as

$$\begin{aligned} R(x) &= E[\theta | y_1 < x] + \bar{\eta} \\ &= \int_{\eta_L}^{\eta_H} g(\eta_1) E[\theta | \theta < x - \eta_1, \eta_1] d\eta_1 + \bar{\eta} \\ &= \int_{\eta_L}^{\eta_H} g(\eta_1) \frac{\int_{\theta_L}^{x-\eta_1} \theta dF(\theta)}{F(x-\eta_1)} d\eta_1 + \bar{\eta} \\ &= \int_{\eta_L}^{\eta_H} g(\eta_1) \left\{ (x - \eta_1) - \frac{\int_{\theta_L}^{x-\eta_1} F(\theta) d\theta}{F(x-\eta_1)} \right\} d\eta_1 + \bar{\eta}. \end{aligned}$$

When $x = \theta_L + \eta_L$, $R(x) = \theta_L + \bar{\eta} > \theta_L + \eta_L = x$.

When $x = \bar{\theta} + \bar{\eta}$, $R(x) = E[\theta | \theta < \bar{\theta} + \bar{\eta} - \eta_1] + \bar{\eta} < \bar{\theta} + \bar{\eta} = x$.

According to the intermediate value theorem, there exists a solution to Equation (21), $y^* \in (\theta_L + \eta_L, \bar{\theta} + \bar{\eta})$.

Since the slope of the left-hand side in Equation (21) is 1, if the slope of the right-hand side is always less than 1 in $(\theta_L + \eta_L, \bar{\theta} + \bar{\eta})$, then y^* is unique.

$$\begin{aligned} \frac{\partial R(x)}{\partial x} &= \int_{\eta_L}^{\eta_H} g(\eta_1) \left\{ 1 + \frac{f(x-\eta_1) \int_{\theta_L}^{x-\eta_1} F(\theta) d\theta}{F(x-\eta_1)^2} - \frac{F(x-\eta_1)}{F(x-\eta_1)} \right\} d\eta_1 \\ &= \int_{\eta_L}^{\eta_H} g(\eta_1) \left\{ \frac{f(x-\eta_1) \int_{\theta_L}^{x-\eta_1} F(\theta) d\theta}{F(x-\eta_1)^2} \right\} d\eta_1 \end{aligned}$$

Following Heckman and Honore (1990), define

$$F_{j+1}(z) = \int_{\theta_L}^{z} F_j(\theta) d\theta, \text{where } F_0 = F(z) \equiv Pr(\theta \leq z).$$

Then

$$\frac{\partial R(x)}{\partial x}=\int_{\eta_L}^{\eta_H} g(\eta_1)\left\{\frac{F_1''(x-\eta_1)F_1(x-\eta_1)}{F_1'(x-\eta_1)^2}\right\}d\eta_1.$$

From log-concavity, $\frac{F_1''(x-\eta_1)F_1(x-\eta_1)}{F_1'(x-\eta_1)^2}\in[0,1]$. A necessary condition for $\frac{F_1''(x-\eta_1)F_1(x-\eta_1)}{F_1'(x-\eta_1)^2}=1$ is $F(x)\propto e^{cx}$, with $c>0$, which is ruled out by the assumption in Subsection 3.1. Therefore, we have

$$\frac{\partial R(x)}{\partial x}<\int_{\eta_L}^{\eta_H} g(\eta_1)\,d\eta_1=1.$$

Thus, y^* is unique.[53]

In the second step, define $\hat{y}=(\hat{y}_1,\hat{y}_2(s),\hat{y}_2(q))$, and $D^3=[\theta_L+\eta_L,\theta_H+\eta_H]\times\left[y^*,\bar{\theta}+\bar{\eta}\right]\times[\theta_L+\eta_L,y^*]$, which is a compact, convex set. Given $\hat{y}_1<\hat{y}_2(s)$, define mappings

$$T(\hat{y})=(T_1(\hat{y}),T_2(\hat{y}),T_3(\hat{y})),$$

where

$$\begin{aligned}
T_1(\hat{y}) &= E[\theta|y_1<\hat{y}_1]+\bar{\eta}+E\{[y_2-\hat{y}_2(q)]\cdot\mathbf{1}(y_2\geq\hat{y}_2(q))|y_1<\hat{y}_1\}\\
&\quad +E[\theta|y_2<\hat{y}_2(q),y_1<\hat{y}_1]-E[\theta|\hat{y}_1\leq y_1<\hat{y}_2(s)],\\
T_2(\hat{y}) &= E[\theta|\hat{y}_1\leq y_1<\hat{y}_2(s)]+\bar{\eta},\\
T_3(\hat{y}) &= E[\theta|y_2<\hat{y}_2(q),y_1<\hat{y}_1]+\bar{\eta}.
\end{aligned}$$

We apply the Brouwer fixed-point theorem to prove there exists a fixed point of $\hat{y}=T(\hat{y})$ in D^3.

It is straightforward to verify the following:

$$\begin{aligned}
T_2(\hat{y}) &\geq E[\theta|\theta_L+\eta_L\leq y_1<y^*]+\bar{\eta}=E[\theta|y_1<y^*]+\bar{\eta}=y^*,\\
T_2(\hat{y}) &\leq E\left[\theta|\theta_H+\eta_H\leq y_1<\bar{\theta}+\bar{\eta}\right]+\bar{\eta}=\left[\bar{\theta}+\bar{\eta}-\bar{\eta}\right]+\bar{\eta}=\bar{\theta}+\bar{\eta}.
\end{aligned}$$

$$\begin{aligned}
T_3(\hat{y}) &\geq E[\theta|y_2<\theta_L+\eta_L,y_1<\theta_L+\eta_L]+\bar{\eta}=\theta_L+\bar{\eta}>\theta_L+\eta_L,\\
T_3(\hat{y}) &\leq E[\theta|y_2<y^*,y_1<\theta_H+\eta_H]+\bar{\eta}\leq E[\theta|y_2<y^*]+\bar{\eta}=y^*.
\end{aligned}$$

$$\begin{aligned}
T_1(\hat{y}) &\geq E[\theta|y_1<\theta_L+\eta_L]+\bar{\eta}+E[\theta|y_2<\theta_L+\eta_L,y_1<\theta_L+\eta_L]\\
&\quad -E\left[\theta|\theta_H+\eta_H\leq y_1<\bar{\theta}+\bar{\eta}\right]\\
&= \theta_L+\bar{\eta}+\theta_L-\left(\bar{\theta}+\bar{\eta}-\bar{\eta}\right)=2\theta_L+\bar{\eta}-\bar{\theta}\\
&\geq \theta_L+\eta_L.
\end{aligned}$$

[53]This y^* is actually the equilibrium threshold value in a two-period asymmetric learning model, where a worker stays with the incumbent firm if and only if $y_1\geq y^*$ at the end of the first period.

$$
\begin{aligned}
T_1(\hat{y}) &\leq E[\theta|y_1 < \theta_H + \eta_H] + \bar{\eta} + E[y_2 - (\theta_L + \eta_L)|y_1 < \theta_H + \eta_H] \\
&\quad + E[\theta|y_2 < y^*, y_1 < \theta_H + \eta_H] - E[\theta|\theta_L + \eta_L \leq y_1 < y^*] \\
&\leq \bar{\theta} + \bar{\eta} + [\bar{\theta} + \bar{\eta} - (\theta_L + \eta_L)] + y^* - y^* \\
&= 2\bar{\theta} + 2\bar{\eta} - (\theta_L + \eta_L) \\
&\leq \theta_H + \eta_H.
\end{aligned}
$$

So $T(\hat{y})$ is a continuous mapping from D^3 to D^3. According to the Brouwer fixed-point theorem, there exists a fixed point which is the solution to equations (15), (19), and (20).

The equilibrium wage settings are proved in the text. ■

C Proof of Proposition 3: Asymmetric versus Symmetric Learning

First, notice that in the symmetric learning model, the worker's turnover decisions in all periods are independent of his ability. Whether he quits the incumbent employer or not depends entirely on the current match quality. During the third period, a worker with ability θ is paid $\theta + \bar{\eta}$ regardless of work history. In contrast, if learning is asymmetric, then we have $W_3^o(s) = \hat{y}_2(s)$ and $W_3^o(q) = \hat{y}_2(q)$. From Equations 19 and 20, we have

$$
\begin{aligned}
W_3^o(s) = \hat{y}_2(s) &= E[\theta|\hat{y}_1 \leq y_1 < \hat{y}_2(s)] + \bar{\eta} \\
&> E[\theta|y_1 < \hat{y}_1] + \bar{\eta} \\
&\geq E[\theta|y_2 < \hat{y}_2(q), y_1 < \hat{y}_1] + \bar{\eta} \\
&= \hat{y}_2(q) = W_3^o(q)
\end{aligned}
$$

Therefore, we have $W_3^o(s) > W_3^o(q)$. ■