ESSAYS ON CAREER DYNAMICS INSIDE ORGANIZATIONS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Xin Jin

August 2014

ESSAYS ON CAREER DYNAMICS INSIDE ORGANIZATIONS

Xin Jin, Ph.D.

Cornell University 2014

This dissertation focuses on the theoretical and empirical analysis of individuals' career and wage dynamics inside firms. It addresses three areas of interests in this field: human capital accumulation, signaling, and organization's structural change.

The first chapter adopts a task-specific human capital perspective to examine the relationship between individuals' horizontal (i.e. lateral movements) and vertical (i.e. promotions) career mobility in a symmetric learning environment. It extends the theoretical literature on individuals' vertical career mobility by incorporating lateral moves in a standard job assignment model with task-specific human capital accumulation. The main intuition is that, when upper-level jobs require a wider set of task skills compared to lower level jobs, firms use lateral moves to develop their employees' task-specific human capital before promoting them. Consequently, lateral moves are positively correlated with individuals' career progressions. This model predicts that individuals who are laterally moved are more likely to be promoted and to experience larger wage growth compared to individuals who do not move. Further, individuals with very high levels of education are less likely to be laterally moved compared to individuals with lower education levels. These predictions are tested using a large employer-employee linked panel on over 30,000 senior managers in more than 500 of the largest U.S. firms during the period of 1981-1985. The empirical evidence supports the theoretical predictions and shows the importance of lateral mobility in individuals' career and wage dynamics.

The second chapter addresses the signaling role of not being promoted and how individuals' wage-profiles are affected by those signals. There is an extensive body of lit-

erature concerning the positive signals associated with promotion. However, theoretical investigations of negative signals associated with non-promotion are nearly nonexistent. In this chapter, a model with asymmetric learning is constructed to capture the negative signals associated with non-promotion. The model shows that, when productivity rises little with additional years on the same job level, the negative signal associated with non-promotion leads to wage decreases. On the other hand, a non-promoted worker's wage increases with additional job-level tenure when additional job-level tenure leads to a sizable increase in productivity. Furthermore, individuals who are promoted when human capital rises little from the previous period earn a lower promotion wage than those who are promoted in a previous period. These predictions are tested using the internal personnel records from a large US firm from 1970-1988. The results support the model's predictions to a large extent. In particular, there is a clear hump-shaped pattern in the wage-job-level-tenure profile for workers who stay in the same job level. This result suggests that, besides determining workers' levels of human capital, job tenure carries rich information about individuals' unobserved ability. The trade-off between negative learning and positive human capital accumulation associated with additional tenure shapes the wage-tenure profile.

The third chapter examine the impact of organizational changes on wages and the wage distribution inside firms. Over the past twenty years, firms became flatter. There is an extensive literature - both theoretical and empirical - that explores the causes of this delayering trend. The consequences of this trend, on the other hand, are not sufficiently studied. This paper examines how wages and the wage distribution change with firm delayering. A job-assignment model with asymmetric information and a slot constraint is considered. The model predicts that more efficient firms are not necessarily larger than less efficient firms if firms are allowed to adjust their internal organizational structure through delayering. After delayering, wages at all levels increase and the wage distribution becomes more unequal. These predictions match a set of empirical findings in recent studies that are not well explained by existing theories.

BIOGRAPHICAL SKETCH

Xin Jin was born in Chengdu, China on January 21, 1984. She received her Bachelor of Engineering from Sichuan University, China in 2006. She began her graduate studies in economics at Cornell University in August 2009. Her primary research focus has been in Labor Economics and Organizational Economics. In September 2014, Dr. Jin will be joining the Department of Economics at the University of South Florida, Tampa as an Assistant Professor.

To my mother and to my husband.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Professor Michael Waldman, the Chairman of my committee, my teacher, advisor, and mentor, for all the hours he spent reading my work and for numerous invaluable insights he gave about my research. His teaching led me to the world of economics, shaped my research perspectives and my view of Economics as a profession. My debt to him goes far beyond my training at Cornell.

I would also like to thank Professor Kevin F. Hallock. Without his encouragement and generous support to grant me access to various data sources, this dissertation would not be possible. I am also grateful to Professors Matthew Freedman and Victoria L. Prowse, the other two members of my committee, for their advice and suggestions about my research and for their encouragement, support, and guidance towards the next stage of my career.

I owe special acknowledgments to Professors Henry Schneider and John Abowd. Professor Schneider gave me his time and insights on numerous occasions in the development of my job market paper and throughout my training at Cornell. Professor Abowd, the Director of Graduate Study, provided with me and my fellow students the timeliest and the most powerful push and support without which this dissertation would not be possible. Special thanks to Professors Michael Gibbs and Wallace Hendricks for sharing their data sources.

I thank Susanne M. Bruyere, Linda Barrington, Sarah von Schrader and all my colleagues at the Institute for Compensation Studies and the Employment and Disability Institute for their support in various research projects. I thank Liyan Yang, Shuang Zhang, Ruoran Gao, and Xiaomeng Lu for their friendship, support, and the time they spent discussing economics with me.

Finally, this thesis would not be possible without the love, encouragement, understanding, and support of my mother and my husband.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

LATERAL MOVES, PROMOTIONS, AND TASK-SPECIFIC HUMAN CAPITAL: THEORY AND EVIDENCE

1.1 Introduction

Most studies on career and wage dynamics inside firms focus on individuals' career movements across different job levels, i.e. promotions (Rosen, 1982; Bernhardt, 1995; Gibbons and Waldman, 1999a). However, empirical and anecdotal evidence suggest that lateral career moves, i.e. individuals' career movements between jobs at the same job level, are very common. For example, Saari et al. (1988) survey 1,000 randomly selected firms in the United States and find that over 40% of the organizations use horizontal moves in their human resource practices. A more prominent example comes from General Electric's most recent succession saga. By the time Jack Welch left the office in 2001, the GE Empire had 12 line departments under the Corporate Executive Office. The current Chairman and CEO, Jeffrey Immelt, served leadership roles in GE Medical, Plastics, and Appliances before becoming the CEO.

These examples show not only that lateral moves are very common but also there is an important link between individuals' lateral moves and promotions. In this essay, I explore how workers' lateral moves relate to their promotions, wage dynamics, and education. To address these questions, I incorporate lateral moves into a job assignment model with task-specific human capital accumulation. The idea of task-specific human capital was first proposed in Gibbons and Waldman (2004). They emphasize that workers' human capital is attached to their jobs. For example, the head of marketing has extensive knowledge about advertisement and consumer behavior, while the head of logistics is an expert on supplier behavior. However, the two managers may have little knowledge about the other's arena although they work in the same firm. That is, employees' knowledge and skills are attached to their jobs and daily routines. This study explores how these task-specific skills play a role in workers' career development. The main argument is that lateral moves help workers to acquire different types of task-specific human capital so that they are more productive when they are promoted. From the theoretical model, I derive a set of predictions regarding the relationship between lateral moves and individuals' career outcomes. I test my model's predictions using a large employer-employee linked panel dataset of U.S. managers and I find empirical support for the model predictions.[1]

To model lateral moves, I consider two job levels in hierarchical firms. There is one job on the upper level (e.g., the Chief Operating Officers) that oversees two lower level jobs (e.g., division managers in marketing and logistics). I allow lateral moves between the two lower level jobs as well as vertical moves across hierarchical levels. The upper-level job uses the task-skills from both of the lower level jobs (e.g., the COO uses both marketing and logistic skills).[2] Those task-skills can be acquired through working in different jobs on the lower level or through formal education. The accumulation of human capital is task-specific and follows diminishing marginal returns to task-tenure in each job.

The market is competitive with free entry. Individuals differ in terms of their innate

[1]Although I provide examples and use data on corporate managers in this study, the lateral-move set up is very general. For example, lateral moves help a cuisine apprentice to become a top chef (Gergaud et. al., 2012).

[2]Surveys and studies that support the assumption that the upper level jobs use a wider set of skills than the lower level jobs are numerous. For example, according to the Occupational Information Network (O*NET), one of the most comprehensive surveys of occupation and job characteristics in the United States, there are 17 core tasks for Chief Executives while there are only five core tasks for Financial Managers at the branch or department level. Mintzberg (1973) in an earlier study shows that managers are presumed to perform a variety of different tasks, while London (1985) argues that horizontal movements are effective to develop managers into "generalists". Ferreira and Sah (2012) provide a rationale for a wider skill set on the upper level. They argue that, because "generalists" can facilitate communication among "specialists", an increasing breadth of expertise with ranks minimizes communication costs.

abilities and their education. An individual's education has two components: a general education level (i.e., years of schooling) and an education type (i.e., majors). When an individual enters the labor market, firms (as well as workers) observe her education level and type but not her innate ability. Firms gradually learn about the individual's innate ability by observing her outputs. The output realization is publicly observable. I further assume that, in the production process, ability and human capital are complementary while education and employment are substitutes.[3]

The firms' problem is to assign workers to jobs optimally so that workers' total expected outputs are maximized. Because the market is competitive, workers' wages are equal to their expected productivity. Therefore, workers have the incentive to accept the job assignment because their total expected wages are maximized when the expected outputs are maximized.

The main trade-off behind the model is the benefit and cost of lateral moves. To see the logic, let us consider the COO and the two division managers in marketing and logistics as an example. Since both marketing and logistics are useful to the COO position, a manager is better off as a COO if she is somewhat knowledgeable about both marketing and logistics. The benefit of lateral moves lies in the fact that the upper-level jobs need a wider but not necessarily deep set of skills. However, not everybody is laterally moved because lateral moves are costly in terms of current productivity because of task-specific human capital. The core idea of task-specific human capital is that individuals cannot fully utilize the skills that she acquires in one job on another. The cost of lateral move is the opportunity cost to the firm and to the worker in terms of

[3]Numerous empirical studies find complementarity between ability and human capital. For example, Bartel and Sicherman (1998) report a positive relationship between training and AFQT scores using the National Longitudinal Survey of Youth. Acemoglu and Pischke (1998) find a similar pattern in German data. In addition, all of my results hold if ability and human capital are not complementary. The assumption that education and employment are substitutes simplifies the model, but none of the results rely on this assumption.

giving up her old skills and to learn new skills in the current period. So, only those individuals who show potential for the COO position are laterally moved. Furthermore, since education and employment are substitutes, when the individual has sufficiently high levels of general education, she does not need to be laterally moved to diversify her skill set because she already has all the skills she needs for the COO position.

My model generates three testable predictions. First, laterally moved individuals are more likely to be promoted compared to individuals who do not move because they have higher expected innate ability and they can reach a higher human capital level upon promotion. Second, laterally moved individuals experience larger wage growth after the moves. Third, the individuals with very high levels of education are less likely to be laterally moved compared to individuals with lower education levels due to the substitution between education and on-the-job human capital development.

To test the model's predictions, I use a unique employer-employee linked panel of over 30,000 senior managers in more than 500 of the largest U.S. firms from 1981 to 1985. Since the seminal work of Baker, Gibbs and Holmstrom (1994a;b), most of the empirical studies on career and wage dynamics, especially the studies that focus on U.S. firms, use a single firm's personnel records.[4] Multi-firm analyses in this area primarily rely on European data (e.g., Devereux et al., 2013; Frederiksen and Kato, 2011). My study uses a multi-firm dataset from the U.S. and generates empirical findings that support my model's predictions. I find positive relationships between lateral moves and promotions and wage growth especially two years after the move, and a negative relationship between lateral moves and years of schooling for more than 19 years of schooling.

This study contributes to the literature in multiple ways. First, it extends the theo-

[4]See Hendricks and Gibbs (2004) for an excellent survey.

retical literature on career and wage dynamics inside firms by formalizing the relationship between workers' lateral moves and subsequent promotions.[5] Second, this study contributes to the human capital literature by exploring the role of task-specific human capital on individuals' career development. It also adds to the discussion on the extent to which formal education and employment are substitutes in individuals' human capital development. Furthermore, this study enriches the empirical literature on career and wage dynamics by providing empirical evidence on a set of new testable implications regarding the relationship between lateral moves and promotions, wage growth, and education using a multi-firm dataset on senior managers in large U.S. corporations.

The outline of the essay is as follows. In Section 2, I review the related literature. Section 3 contains the analysis of a three-period model under full information and a model with symmetric learning. In Section 4, I test the model's predictions and discuss the empirical findings. I discuss alternative explanations in Section 5. Section 6 provides concluding remarks with a discussion of some possible extensions of this study.

1.2 Related Literature

Most of the theoretical literature on individuals' career dynamics inside firms focuses on how individuals move along the job ladder vertically. Two of the building-block theoretical models concerning (vertical) career movements are the tournament model (e.g., Lazear and Rosen, 1981; Rosen, 1986) and the job assignment model (e.g., Gibbons and Waldman, 1999a). Both types of models assume only one type of job at each level, and lateral moves are not considered.

[5] Gibbons and Waldman (2004) touched on this relationship very briefly. Conaty and Charan' in their 2010 New York Time's Bestseller book, Talent Matters, has some discussion about using job rotation as a training device. This study formalizes and extends their observations and make the underlying mechanism transparent.

A few studies have considered two specific types of horizontal movements: job rotation and horizontal transfers that include promotions. The first type of horizontal move is job rotation. Job rotation is a special type of lateral move in which each trainee in the job rotation program follows a predetermined career path by changing positions within the same job level (e.g., medical or management trainees in their training period usually rotate around different departments). Ortega (2001) considers a model where job rotation facilitates the firms' learning about their workers' ability match with different jobs. Building on Ortega's idea, Li and Tian (2013) investigate a directed search model where job rotation improves the match between workers and jobs and thus leads to higher wages and lower turnover rates in larger firms. However, neither of these two job-rotation models makes predictions regarding the relation between horizontal moves and individuals' career progressions.

The other type of horizontal movement investigated in the literature is horizontal transfers that include promotions (Kusunoki and Numagami, 1998; Ariga, 2006; Sasaki et. al., 2012). Horizontal transfers defined in those studies differ from lateral moves that defined here because those transfers are not restricted to movements within the same job level. One prediction from this literature is that a current period transfer is more likely to happen with a current period promotion, i.e., workers are transferred when they are promoted (Sasaki et. al., 2012). However, these models do not make predictions about current transfers and subsequent promotions as my model does.[6]

My study also relates to the human capital literature. In particular, I adopt the task-specific human capital approach developed in Gibbons and Waldman (2004; 2006) rather than the traditional Beckerian dichotomy of general and firm-specific human cap-

[6]Friebel and Raith (2013) also consider a model with cross-divisional transfers. The focus of their paper is on the agency problem and job assignment efficiency under contract arrangements with and without horizontal transfers. They show that the contract with horizontal transfers dominates the contract without when managers have private informationn about their workers in their own division and when the division profit is more sensitive to division output (thus correct job assignment is more important).

ital (Becker, 1962;1964).[7] There is substantial evidence that supports the importance of task-specific human capital in workers' wage dynamics. Using German data, Gathmann and Schonberg (2010) empirically quantify the significant contribution of task-specific human capital to wage growth. They find that task-specific human capital accounts for 22% to 52% of individuals' overall wage growth. Using a sample of 1% of the British workforce, Devereux et al. (2013) find that a large proportion of the return-to-tenure arises with job-level tenure within firms rather than firm-level tenure. Using data from 76 firms in the U.S. Information Technology industry, Schulz et al. (2013) show that task-specific human capital (measured as job-tenure) is positively associated with employee compensation. My study goes one step further by exploring the way in which task-specific human capital is acquired through lateral moves.[8]

There are two studies on lateral moves that are closely related to my analysis. Gittings (2012) uses the same data set that I use and shows that individuals who are laterally moved have higher wages in the period of the lateral move. However, he does not provide a theoretical explanation for this empirical finding. Clemens (2012) develops a theoretical prediction that lateral transfers into a fast job is positively correlated with subsequent promotions but transfers into a slow job is negatively correlated with subsequent promotions.[9] However, using a single firm's personnel records, Clemens finds that lateral transfers into either fast or slow jobs are positively correlated with subsequent promotions. That is, he does not find empirical support for his model's prediction but

[7] Other studies using Gibbons and Waldman's (2004;2006) framework concerning task-specific human capital accumulation include Balmaceda (2006) and Clemens (2012). Balmaceda (2006) applies the task-framework to consider optimal job designs. Clemens' (2012) analysis is closely related to my essay and is discussed further below.

[8] The idea that human capital is attached to jobs is also closely related to the occupational-specific human capital (Kambourov, et al., 2009) and industry-specific human capital (Parent, 2000) approaches. Both approaches find supporting evidence that human capital does not accumulate homogeneously and occupational and industry-specific human capital are more pertinent to wage profiles rather than firm-specific human capital.

[9] Clemens defines fast jobs as those positions in a firm out of which promotions are more likely compared to other positions at the same hierarchical level. The "other" positions with a lower rate of promotion are referred to as slow jobs.

provides some evidence that is consistent with my model's prediction regarding lateral moves.

In summary, this study extends the theory on job-assignment and the literature on human capital development. It also enriches the empirical studies on lateral moves by providing a theoretical framework and a set of new testable implications.

1.3 Theoretical Analysis

In this section, I set up a three-period two-level model to capture lateral moves and career progression. The production technology is closely related to those analyzed in Gibbons and Waldman (2006). I first consider a model with full information. I then develop a model with symmetric learning and derive testable implications.

1.3.1 The modeling environment

There is free entry into production. All firms are identical and labor is the only input. Workers and firms are risk-neutral and they do not discount the future. Workers bear no cost to change firms and firms bear no cost to hire or fire workers. Workers and firms enter into the employment relation through spot-market contracting.

Each firm consists of two hierarchical levels. There are two jobs on level 1, denoted by $j,\ j \in \{A,B\}$, and only one job on level 2. Workers' careers last for three periods, denoted by $t \in \{1,2,3\}$. Workers can move laterally between lower level jobs and vertically across job levels. Workers in period t have $t-1$ periods of labor market experience, which is equal to the sum of job tenure in each job at each level.

Worker i enters the labor market in period 1 with a general education level s_i, where s_i takes integer values from 1 through N, i.e. $s_i \in \{1, \ldots N\}$. Workers also differ by their education types. There are two education types: type A and type B. There are $2 \cdot m$ workers in each education level and half of them have type-A education.[10] Within each education type at each education level, workers' innate abilities are random draws from a common distribution such that a worker has high innate ability θ_H with probability p_0, and low innate ability θ_L with probability $1 - p_0$.[11]

A particular education type gives a worker α units of match quality in the corresponding job type. Let α_{ij} denote worker i's match with job j. If worker i has type-A education, then $\alpha_{iA} = \alpha, \alpha_{iB} = 0$. If worker i has type-B education, then $\alpha_{iA} = 0, \alpha_{iB} = \alpha$.[12] The education type can be understood as a worker's major or curriculum focus in school that makes her a better fit with one job or the other. A worker with type-j education is referred to as a type-j worker. Job j is referred to as type-j workers' matched job. Like education levels, individuals' education types are fully observable to all labor market participants. I further assume that if a worker is not assigned to her matched job in her first employment period, the match quality depreciates and only $\lambda\alpha(\lambda \in [0,1])$ is applicable in the next period. In addition, this match quality is not directly applicable to the level-2 job. If a worker has never worked in her matched job on the lower level, she cannot apply this match on the upper-level job.

Let $\tilde{x}_{ijt} = x_{ijt} + \alpha_{ij}$ denote worker i's *effective* job-specific task-tenure in job j on level 1 in period t, where x_{ijt} is worker i's tenure in job j on level 1 prior to t and α_{ij} is

[10] The proportion of type-A education is not essential for the result.

[11] Note that in this set up, workers' education levels are not affected by their innate abilities. Alternatively, I can assume that there are more high ability workers in the high education groups than in the low education groups. But this assumption would not change any of the results concerning lateral moves and promotions.

[12] To keep the model tractable, I assume a perfect match between education types and job types. I also assume a constant match quality. By doing so, I abstract away the effect of lateral moves on improving workers' matches with jobs and focus on the human capital development effects from lateral moves.

worker i's match with job j. x_{i2t} is worker i's tenure on level 2 prior to t. The job experience, education level and match quality add to workers' stock of human capital. Greater experience, higher education level, and a better match of education types to jobs make a worker more productive. I consider the set-up where schooling and job experience are substitutes (Mincer, 1958;1962) while ability and human capital are complements (Acemoglu and Pischke, 1998). The production technology in job j on level 1 is

$$y_{ijt} = d_1 + c_1[\theta_i f(\tilde{x}_{ijt} + s_i) + \varepsilon_{ijt}],\ j \in \{A, B\}. \tag{1.1}$$

$f(\cdot)$ captures individuals' human capital accumulation.$\theta_i f^{'}(\cdot)$ is the speed with which human capital grows in period t. Following Acemoglu and Pischke (1998), I assume that $f(\cdot)$ is twice continuously differentiable, strictly increasing and concave with $f(0) > 0$. I also assume that $lim_{\tau \to \bar{\tau}} f^{'}(\tau) = 0$ for some $0 < \bar{\tau} < N$, where N is the highest level of general schooling. This restriction says that a worker cannot learn more from work after she spends sufficient amount of time on the job or if she has a very high level of general schooling (such as a Ph.D).

In my model, human capital acquisition is task-specific in two ways. First, the task-skills acquired in one of the lower level jobs are not applicable to the other job. Second, a proportion ($1-\gamma$, $0 \leq \gamma \leq 1$) of the task-skills are lost when a worker is promoted from the lower level jobs to the upper-level job. Formally, for workers with $\tilde{x}_{iAt}$ effective task-tenure from job A, $\tilde{x}_{iBt}$ effective task-tenure from job B, and x_{i2t} task-tenure from level 2, her output in the level-2 job in period t is

$$y_{i2t} = d_2 + c_2\left\{\theta_i[f(x_{i2t} + \gamma\tilde{x}_{iAt} + s_i) + f(x_{i2t} + \gamma\tilde{x}_{iBt} + s_i)] + \varepsilon_{i2t}\right\}, \gamma \in [0,1]. \tag{1.2}$$

Equation (1.2) tells us that the upper-level job uses task-skills from both jobs on the lower level. Given the concavity in $f(\cdot)$, having extensive experience in one of the lower level jobs provides little incremental total human capital as opposed to having a more balanced skill set in both jobs.[13] I assume the general education level is fully applicable to all jobs across levels (as opposed to the special component of a worker's education, α_{ij}, which is only applicable to the matched job). I further assume that none of the task-tenure on the upper-level job can be applied to the lower level jobs. When a worker is demoted, she loses all the task-tenure on level 2.

d_L and c_L, $L \in \{1,2\}$, are production constants known to all labor-market participants. I assume that $d_1 > d_2$ and $c_1 < c_2$, which means output increases faster with ability in the upper level job (Rosen,1982 ; Waldman, 1984a). ε_{ijt} and ε_{i2t} are noise terms drawn from a normal distribution with mean 0 and variance σ_ε^2.

I focus on Bayes-Nash Equilibria. The timing of the events in the economy is the following. At the beginning of period 1, firms observe an individual's education level and education type but not her ability type. Firms make job assignment decisions and wage offers based on the individual's expected abilities, education levels and education types. The individual then chooses the firm that offers the highest wage to work at. At the end of period 1, all firms observe each worker's output and update beliefs about abilities. At the beginning of period 2, firms make job assignment decisions and wage offers based on workers' expected abilities, education levels and education types. Firms can either promote a worker to level 2, move a worker to a different job, or let her stay in the same position. Workers then choose the firm that offers the highest wage to work at. If multiple firms offer the same highest wage, a worker randomly chooses among those firms but stays with her period-1 employer if the period-1 employer is one of the highest-wage-offer firms. Production begins. At the end of period 2, workers' outputs are observed. This process is repeated in period 3.

[13] Also note that the human capital accumulation functions on level 1 and level 2 do not have to be the same. I assume they are the same to simplify the parameter restrictions.

1.3.2 Analysis and Testable Implications

In this subsection I start by describing equilibrium behavior in a benchmark model where individuals' abilities are fully observable. I then consider a model with symmetric learning.

When task-specific human capital accumulation is assumed, strategic firms do not simply maximize current productivity since current job assignment decisions (i.e., promotions and lateral moves) affect future human capital development. Instead, firms choose the optimal career development path in a given period that maximizes workers' total expected productivity over the remaining periods. In addition, firms make zero expected profits in equilibrium in each period due to free entry and the absence of firm-specific human capital. Since there is no firm-specific human capital, workers' wages are equal to their expected productivity in each period.

Equilibrium with Full Information

I start the analysis by considering all possible career paths under full information. To reduce the number of cases, I focus on parameterizations for which the most able worker with the highest education level is not assigned to level-2 in period 1.[14] Furthermore, I focus on the parameterization such that if a worker is not assigned to her matched job in the first period, a substantial amount of the match quality is lost.[15] Under these two restrictions, a worker is always assigned to the job that matches with her education type (or major) when she first enters the labor market. In period 2, a worker can be promoted, laterally moved or held in the same position. In period 3, since a demoted worker loses

[14] The parameter restriction that guarantees this is given in the Appendix.

[15] Recall that if a type-j worker is not assigned to job-j in her first employment period, only $\lambda\alpha$ of the match quality is applicable in the next period. Here I focus on the equilibrium where λ is sufficiently small.

all her task-tenure on the upper level, it is not efficient to first promote then demote a worker. Also, under full information, firms have no incentive to laterally move a worker in period 2 but not promote her in period 3.

Based on the above analysis, there are four potential equilibrium paths: starting in their matched job in period 1, workers can be (i) promoted in period 2 and stay on level 2 in period 3; (ii) held in their matched jobs in period 2 and promoted in period 3; (iii) laterally moved in period 2 and promoted in period 3; (iv) stay in their matched jobs for all three periods. Since all workers are assigned to their matched jobs in period 1, I only need to consider the optimal output from period 2's perspective. It can be shown that for all positive values of ability types, either career path (ii) dominates career path (iii) or verse visa. When stay-promote (career path (ii)) dominates lateral-move-promote (career path (iii)), there are no lateral moves in equilibrium. I refer to parameterizations that yield this equilibrium outcome as the non-lateral-move regime. Similarly, when lateral-move-promote (career path (iii)) dominates stay-promote (career path (ii)), lateral moves exist in equilibrium. I refer to parameterizations that yield this equilibrium outcome as the lateral-move regime. The parameterizations that sustain the lateral-move regime satisfy equation (1.3).

$$c_2[f(\gamma+\gamma\alpha+s)+f(\gamma+s)-f(2\gamma+\gamma\alpha+s)-f(s)] > c_1[f(1+\alpha+s)-f(s)], s < \bar{\tau} < N \tag{1.3}$$

Condition (1.3) guarantees that, given a low education level s, lateral moves lead to a larger gain in human capital upon promotion (LHS) than what is lost upon moving (RHS).

In the lateral-move regime, the very high ability workers are promoted in period 2. If a worker is not good enough for an immediate promotion, she is laterally moved in preparation for a promotion in period 3. Proposition 1 summarizes equilibrium behavior in the lateral-move regime. Let w_{it} denote the wage paid to worker i in period t. All proofs are given in the Appendix.

Proposition 1. *(full-information model) If each worker's innate ability is fully observable and equation (1.3) is satisfied, then there exist two critical values, $\tilde{\theta}^*$ and θ^*, $\tilde{\theta}^* < \theta^*$, such that job assignment rules and wages are given by (i) through (iv):*

(i) If $\theta_i > \theta^*$, then worker i is promoted to level 2 in period 2 and $w_{i2} = d_2 + c_2\theta_i[f(\gamma+\gamma\alpha+s)+f(s)]$. This worker remains on level 2 in period 3 and $w_{i3} = d_2 + c_2\theta_i[f(1+\gamma+\gamma\alpha+s)+f(1+s)]$.

(ii) If $\tilde{\theta}^* < \theta_i \leq \theta^*$, then worker i is laterally moved in period 2 and $w_{i2} = d_1 + c_1\theta_i f(s)$. This worker will be assigned to level 2 in period 3 and $w_{i3} = d_2 + c_2\theta_i[f(\gamma+\gamma\alpha+s)+f(\gamma+s)]$.

(iii) If $\theta_i \leq \tilde{\theta}^*$, then worker i remains in her matched job in periods 2 and 3 with $w_{i2} = d_1 + c_1\theta_i f(1+\alpha+s)$ and $w_{i3} = d_1 + c_1\theta_i f(2+\alpha+s)$.

(iv) In period 1, all workers are assigned to their matched jobs on level 1 and $w_{i1} = d_1 + c_1\theta_i f(\alpha+s)$.

Now consider how lateral moves are related to promotions and wage changes in the lateral-move regime. Proposition 1 indicates that, under full information, a worker is promoted in period 3 if and only if she is laterally moved in period 2 (given she is not promoted in period 2). The reason is that, with perfect information, firms have no incentive to incur the cost to move a worker if they do not expect to promote this worker, since the gain of a lateral move is only realized upon promotion. Thus, under full information, lateral moves in period 2 predict promotion in period 3 with probability 1.

With regard to wage changes, workers who are laterally moved have larger wage growth in period 3 relative to workers who are not moved (and are not promoted in period 2). Two forces lead to this result. First, laterally moved workers have higher innate abilities than non-movers. Second, lateral moves lead to a larger increase in human capital upon promotion. To understand the logic, note that the increase in human capital for the non-movers comes from one additional period of tenure in her matched job on level 1. The increase in human capital for the movers comes from the difference between the human capital level she achieves on level 2 and the human capital level in her non-matched job on level 1. Since the movers achieve a higher human capital level after the move but their human capital level in the period of lateral moves is lower than the non-movers, the movers have a larger human capital growth compared to the non-movers. Consequently, laterally moved workers experience a larger wage growth than non-movers in periods after the move.[16]

Now consider how education affects the equilibrium. Let us first consider the equilibrium when the general education level is sufficiently high, in particular when $s_i \geq \bar{\tau}$ (recall that $lim_{\tau \to \bar{\tau}} f^{'}(\tau) = 0$ for some $0 < \bar{\tau} < \infty$). In this case, a worker is almost equally competent in both jobs and she can learn very little from the other job through a lateral move. Thus, the worker enters into the non-lateral move regime, i.e., equation (1.3) is not satisfied. In period 2, she is either promoted or stays in her incumbent job for one more period before promotion in period 3. Compared to a worker whose general education is not that high (i.e., $s_i < \bar{\tau}$) so she can reach a higher human capital level from a lateral move (i.e., equation (1.3) is satisfied), the worker with a high education

[16] In the current set up, a laterally moved worker has a wage decrease in the period of the lateral move under full information. This is because I only consider on-the-job task-specific human capital accumulation in the model. If general human capital is included, a laterally moved worker may have a wage increase in the period of the lateral move. In addition, if general human capital is complementary to workers' innate abilities, the mover may have a larger wage growth than a non-mover in the period of the lateral move even though there is a task-specific human captial loss to the mover. I will come back to this point when I discuss the model with symmetric learning.

level is less likely to be laterally moved.

Next, consider the equilibrium when the general education level is not high enough to fully substitute for the human capital gain from lateral moves (i.e. $s_i < \bar{\tau}$) and the return to lateral moves is high upon promotion (i.e., equation (1.3) is satisfied). In this regime, the net benefit of being laterally moved increases with the general education level when the upper-level job is substantially different from the lower level jobs (i.e. $c_2 \gg c_1$). This is because the net benefit from moving a higher-education worker is higher than that of moving a lower-education worker. However, when the upper-level job is not substantially different from the lower level jobs, even in the lateral move regime, individuals with more education can be less likely to be laterally moved due to the substitution between education and task-tenure.

Equilibrium with Symmetric Learning

The full-information model carries most of the insights regarding how lateral moves affect promotion probabilities, but the prediction that lateral moves predict promotions perfectly is not realistic. In this subsection, I consider equilibrium behavior with symmetric learning. For tractability, I focus on parameterizations that satisfy (1.3) in the main part of the analysis and relax this restriction when I consider the relationship between education and lateral moves.

With symmetric learning, all firms have the same information about a worker's innate ability. Define the information that each firm gets in period t as $z_{it} = (y_{ijt} - d_l)/c_l$. Let θ^e_{it} denote the expected ability in period t conditional on the information. That is, $\theta^e_{it} = E(\theta|z^t)$ where $z^t = \{z_{t-1}, z_{t-2}, \ldots, z_1\}$ is the full history of information.

I consider the job assignment problem with symmetric learning by backward induction. In period 3, since it is the last period, if a worker is not promoted, it is optimal to assign the worker to her matched job. That is, if workers are not promoted in period 3, a type-A worker will be assigned to job A while a type-B worker will be assigned to job B, regardless of their assignments in period 2. Also, since period 3 is the last period, job assignments in period 3 are free of investment concerns. Thus, the equilibrium job assignments maximize period 3's productivity. Let θ_3^{jk} and θ_3^{jj} denote, respectively, the cutoff ability levels for promotion in period 3 for a period-2 mover and a period-2 non-mover (a non-mover is not promoted or moved). θ_3^{jk} solves $d_1 + c_1\theta_3^{jk} f(1+\alpha+s) = d_2 + c_2\theta_3^{jk}[f(\gamma+\gamma\alpha+s) + f(\gamma+1+s)]$, while θ_3^{jj} solves $d_1 + c_1\theta_3^{jj} f(2+\alpha+s) = d_2 + c_2\theta_3^{jj}[f(2\gamma+\gamma\alpha+s) + f(s)]$. Let θ_3^{j2} denote the cutoff ability level for a period-2 promoted worker to stay on level 2 in period 3. θ_3^{j2} solves $d_1 + c_1\theta_3^{j2} f(1+\alpha+s) = d_2 + c_2\theta_3^{j2}[f(1+\gamma+\gamma\alpha+s) + f(1+s)]$.

In period 2, firms choose job assignments to maximize workers' total expected outputs in periods 2 and 3. Let θ_2^j denote the cutoff ability level for promotions in period 2. θ_2^j solves $d_1 + c_1\theta_2^j f(s) + E(y_{i3}|\theta_{i2}^e \leq \theta_2^j) = d_2 + c_2\theta_2^j[f(\gamma+\gamma\alpha+s) + f(s)] + E(y_{i3}|\theta_{i2}^e > \theta_2^j)$. $E(y_{i3}|\cdot)$ is the expected output in the next period given this period's expected ability and job assignment. Let $\tilde{\theta}_2^j$ denote the cutoff ability level for lateral moves in period 2. $\tilde{\theta}_2^j$ solves $d_1 + c_1\tilde{\theta}_2^j f(1+\alpha+s) + E(y_{i3}|\theta_{i2}^e \leq \tilde{\theta}_2^j) = d_1 + c_1\tilde{\theta}_2^j f(s) + E(y_{i3}|\theta_{i2}^e > \tilde{\theta}_2^j)$. The task-tenure in the incumbent job and the match quality are unutilized in the period when a worker is moved to a new job.

Similar to the full-information case, the optimal job assignment rules under symmetric learning are characterized by cutoff ability levels. The difference is that, with production uncertainties, firms sometimes "make mistakes". In the full-information case, only period-2 lateral movers are promoted in period 3. Under symmetric learning, if a non-mover produces very high output in period 2, she might be promoted in period

3. On the other hand, if a mover produces very low output, she might be promoted in period 3. Therefore, there are two types of workers at the beginning of period 3: those who were laterally moved in period 2 and those who were not laterally moved, and both of them can be considered for promotions.

Proposition 2 summarizes the above discussion.

Proposition 2. *(symmetric learning) Suppose that learning is symmetric and the prior belief about a worker's type is that a worker has high innate ability θ_H with probability p_0 , and low innate ability θ_L with probability $(1-p_0)$. The job assignment rules and wages are given by (i) through (iv):*

(i) In period 3, if worker i's expected ability $\theta^e_{i3} > \theta^{j2}_3, j \in \{A,B\}$ and she was promoted in period 2, then she remains on level 2. If worker i's expected ability $\theta^e_{i3} > \theta^{jk}_3($ or $\theta^{jj}_3)$ and she was (or was not) laterally moved in period 2, then she is assigned to level 2. All non-promoted workers in period 3 are assigned to their matched jobs on level 1.

(ii) In period 2, if worker i's expected ability $\theta^e_{i2} > \theta^j_2, j \in \{A,B\}$, then she is promoted to level 2; if $\tilde{\theta}^j_2 < \theta^e_{i2} \leq \theta^j_2$, then she is laterally moved; if $\theta^e_{i2} \leq \tilde{\theta}^j_2$, then she remains in job-j on level 1.

(iii) In period 1, all workers are assigned to their matched jobs on level 1 and $w_{i1} = d_1 + c_1\bar{\theta} f(\alpha+s)$, $\bar{\theta} = p_0\theta_H + (1-p_0)\theta_L$.

(iv) $\theta^{jk}_3 < \theta^{jj}_3$. The cutoff ability for promotions in period 3 is lower for lateral movers.

I now derive two testable predictions from the equilibrium described in Proposition 2. The first prediction considers the relationship between lateral moves and expected promotion probabilities. As discussed above, lateral moves incur an opportunity cost in human capital in period 2. However, if a period-2 lateral mover is promoted in period 3, she can achieve a higher human capital level than a non-mover. Since the lateral-move decisions are endogenous, firms would want to choose a cutoff level such that the expected probability that a laterally moved worker is promoted in the next period is high enough to ensure that the expected return from each lateral move exceeds the expected cost in task-specific human capital upon moving. Thus, from period 2's perspective, the probability that a laterally moved worker is promoted in period 3 will be higher than the probability that a non-mover is promoted. Formally, let $X = prob.(\theta_{i3}^{e} > \theta_{3}^{jk} | \tilde{\theta}_{2}^{j} < \theta_{i2}^{e} \leq \theta_{2}^{j})$ denote the probability that a laterally moved worker is promoted in period 3 from period 2's perspective. It is the probability that worker i's expected ability in period 3 conditional on her output history $\{z_{i1}, z_{i2}\}$ is greater than the promotion threshold for movers (θ_{3}^{jk}), given that worker i is laterally moved in period 2 ($\tilde{\theta}_{2}^{j} < \theta_{i2}^{e} \leq \theta_{2}^{j}$). Similarly, $Y = prob.(\theta_{i3}^{e} > \theta_{3}^{jj} | \theta_{i2}^{e} < \tilde{\theta}_{2}^{j})$ is the probability that a non-mover is promoted in period 3.

Corollary 1. *Under symmetric learning, the probability that a period-2 laterally moved worker is promoted in period 3 is larger than the probability that a period-2 non-laterally moved worker is promoted in period 3 (given she is not laterally moved and not promoted in period 2), i.e.* $X > Y$.

The positive relation between lateral moves and subsequent promotions come from two sources. First, from (iii) in Proposition 2, laterally moved workers have higher expected ability than non-movers. Second, from (iv), the movers have a lower threshold to pass in order to get the promotion. That is, the movers are more likely to be promoted both because they are better workers and because they are on a better path to accumulate human capital for the upper level job.

The second testable implication concerns the size of the wage changes in periods after lateral moves. The symmetric learning model generates similar predictions regarding wages as those in the full-information model, i.e., laterally moved workers are expected to have a larger wage growth in period 3. Let Δw^L and Δw^S denote expected wage changes (in levels) for lateral movers and non-movers from period 2 to period 3.

Corollary 2. *Conditional on prior beliefs about workers' abilities, laterally moved workers on average experience a larger wage growth in the period after the move, i.e.* $\Delta w^L > \Delta w^S$.

As in the full information model, the positive relationship between wage growth and lateral move comes from positive learning about a laterally moved worker's innate ability and the effect of a lateral move on human capital accumulation. The positive learning on lateral moves follows exactly the same logic as in the full information model. The relationship between lateral moves and the expected human capital level is the following. As stated in Corollary 1, lateral moves are associated with a higher promotion probability. In addition, a laterally moved worker can achieve a higher human capital level upon promotion. Thus the mover's expected human capital level is higher than a non-mover's human capital level. Since in equilibrium workers' wages are equal to their expected productivity, laterally moved workers are expected to have a larger wage growth. This positive relationship between lateral moves and wage growth exists even when ability types are controlled for since lateral moves have a permanent effect on human capital accumulation.

Now consider wage growth in the period of a lateral move. Under the current set up, the model makes an ambiguous prediction. The reason is that although lateral movers have a decrease in their (task-specific) human capital levels in the period of the lateral move, there is positive learning about the movers' expected abilities. If the positive learning about lateral movers' abilities outweighs the decrease in human capital, movers experience a larger wage growth in the period of lateral moves compared to non-movers. If the decrease in human capital outweighs the positive learning on ability, movers have a smaller wage growth in the period of lateral moves compared to non-movers.

The third prediction is not directly from Proposition 2. It considers the relationship between lateral moves and general education levels. As discussed in the full information case, when the general education level is sufficiently high, workers enter the non-lateral move regime.[17] However, for workers with lower education levels such that the lateral-move regime is sustained, the probability of lateral move is positive in equilibrium. Therefore, workers with very high education are less likely to be laterally moved.[18] I summarize this argument in Corollary 3.

Corollary 3. *In period 2, among those who are not promoted in the current period, the probability of being laterally moved decreases with the general education when the schooling level is sufficiently high.*

Now consider the relationship between lateral moves and education levels in the lateral move regime. It can be shown that if the upper level job is sufficiently different from the lower level job (i.e., $c_2 \gg c_1$), individuals with higher education levels are more likely to be laterally moved. However, since I cannot measure the difference in the production functions in the two jobs levels, the empirical prediction regarding the probability of lateral moves and education levels is ambiguous for workers with low and medium education levels.

In this section, I incorporate lateral moves into a model with job assignment, human

[17]This suggests that there are no lateral moves with very high education. I observe this pattern in the data. I will return to this point in the empirical section.

[18]Note that, if the match quality (α_{ij}) is very high, it would be too costly to laterally move someone. Therefore, there would be no lateral moves in the equilibrium. Empirically, we cannot perfectly distinguish between the effect from education-tak-tenure substitution and the effect from match cost.

capital development, and symmetric learning. Since lateral move is a more efficient way (relative to no lateral moves) to develop human capital for the upper level job, relatively higher ability workers are selected into this career path. Consequently, lateral moves are associated with positive career outcomes in terms of promotion probabilities and wage growth in periods after the move. Furthermore, due to the substitution between education and human capital development on-the-job, workers with very high levels of education are less likely to be laterally moved.

1.4 Data and Tests

In this section, I test my model's predictions using a large employer-employee linked panel on top American executives. Data on top managers are suitable for this study for two reasons. First, the nature of the management level jobs matches the assumption that the upper-level jobs use a wider set of skills than the lower level jobs. For example, it is important for a top-level executive (such as a Chief Operations Officer) to be familiar with business operations in different lower-level line departments, but a research associate in the R&D department might not need task-skills from the Finance department. Second, lateral moves (as defined in my essay) are more likely to exist among senior managers. Campion et al. (1994) find that horizontal movements among upper level managers bring "broader perspective on other business functions" to the managers, while young workers are more likely to be involved in job rotation.

In the following subsections, I first describe the data used for this study and how the sample and measures are constructed. I then test the three predictions derived in the previous section.

1.4.1 Data

The data is a large employer-employee linked panel that contains information on over 30,000 executives in over 500 of the largest U.S. firms during the period 1981-1985.[19] A unique identifier is assigned to each firm. However, the same individual may have different identifiers in different firms, which means I cannot track individuals across firms.

There are three compensation-related variables (in nominal terms): base pay, bonus pay, and pay midgrade. I deflate them using the Consumer Price Index in 1982 US dollars provided by the Bureau of Labor Statistics. I construct the measure of total pay as the sum of base pay and bonus pay.

Three variables in this dataset define an executive's position in the firm: reporting level, organizational unit level, and job code. The reporting level counts the number of levels away from the Board of Director (the BOD). The CEO directly reports to the BOD and thus is at reporting-level 1. All executives who directly report to the CEO are at reporting-level 2, and all executives who directly report to level-2 executives are at reporting-level 3, etc. The organizational unit level counts the major organizational units away from the BOD. An organizational unit is a company, group, division, sales region, or manufacturing facility that the company accounts for as a separate profit center. In a hypothetical organization where a division manager reports to a group executive who reports to a corporate executive, the division manager is on Unit Level 3, the group executive is on Unit Level 2, and the corporate executive is on Unit Level 1. Job codes

[19]The dataset was constructed by a large consulting firm through annual surveys of those firms. Firms are paid to participate in the survey. Each firm reports data on about 80 executives per year. The dataset contains rich information on individual, job and firm characteristics, including: age, years of education, hiring date, job title, reporting level, unit level, job title, base pay, bonus pay, pay midgrade, firms' industry, profits, sales, span of control, and employment size. See Abowd (1990), Bognanno(2001), Belzil and Bognanno (2008), Gittings (2012) and Belzil et al. (2012) for more details about the data and the data collection procedure.

are workers' job titles. There are 11 reporting levels, 8 unit levels, and 165 job codes (i.e. job titles).

To define job transitions, I use the reporting level as a basic measure. I define promotion as an upward movement in the reporting level (e.g., from level 4 to level 3). Since executives in different unit levels can share the same job title, I do not restrict promotions to upward movements with a job title change.[20] However, I define demotion as a downward movement in the reporting level with a job title change. As Belzil et al. (2012) point out, there is an organizational re-structuring through the sampling years from 1981 to 1985 where a COO position is added between CEO and lower level executives.[21] This organizational change causes a universal downward movement of reporting levels without actual demotions or job title changes. Therefore, to qualify for a demotion in the data, an individual has to move down the reporting level with a job title change.

Defining lateral moves requires more careful work. Conceptually, a lateral move entails a change of job content without a change in hierarchical position within a firm. Due to data constraints, I cannot identify lateral moves between different divisions within the same unit level. Therefore, I define lateral moves as movements within the same reporting level with a job title change or a unit-level change. These types of lateral moves provide executives with the opportunities to acquire different types of task-specific skills in another job or another business unit.[22]

Table 1.1 summarizes different types of lateral moves and subsequent promotions

[20]For example, a transition from the Top Personnel Executive in a profit center to the Top Corporate Personnel Executive comes with no job title change but is clearly a promotion.

[21]In a recent study, Caliendo et al. (2013) find in a comprehensive dataset of French manufacturing firms that firms frequently expand by adding layers (levels).

[22]As a robustness check, I exclude those lateral moves with a unit level up but without reporting level changes since this type of movement can be considered as a promotion. The results are consistent with what I find with the definition of lateral move used in the main part of the analysis.

Table 1.1: Lateral Moves and Subsequent Promotions by Type of Moves.

	Promotion in One Year		Promotion in Two Years	
	No	Yes	No	Yes
Lateral Move				
Mover	3,314	533	1,524	250
	86.2%	13.9%	85.9%	14.1%
Non-mover	25,681	3,675	12,471	1,752
	87.5%	12.5%	87.7%	12.3%
Lateral Move with Unit-level Up				
Mover	1,051	151	415	87
	87.4%	12.6%	82.7%	17.3%
Non-mover	27,944	4,057	13,580	1,915
	87.3%	12.7%	87.6%	12.4%
Lateral Move with Unit-level Down				
Mover	888	185	398	81
	82.8%	17.2%	83.1%	16.9%
Non-mover	28,107	4,023	13,597	1,921
	87.5%	12.5%	87.6%	12.4%
Lateral Move with Unit-level Same				
Mover	1,375	197	711	82
	87.5%	12.5%	89.7%	10.3%
Non-mover	27,620	4,011	13,284	1,920
	87.3%	12.7%	87.4%	12.6%

Note: Sample restricted to executives who are not promoted in the current period.

one year and two years after the move. Overall, non-promoted executives with any kind of lateral move are more likely to have a promotion one year or two years after the move than non-promoted executives without lateral moves. For example, among those who are laterally moved, 13.9% are promoted one year after the move, while only 12.5% of the non-movers are promoted one year from the period of lateral moves. The gap in the promotion probability becomes larger two years after the lateral-move period. In the second year after the moves, 14.1% of the movers are promoted while 12.3% of the non-movers are promoted. In the following analysis, I consider any lateral moves within the same reporting level either with a job-title change or with a unit-level change. I further restrict the sample to executives who appear in the sample for at least 3 consecutive

Table 1.2: Summary Statistics.

A. Excecutive Characteristics in 1981					
No. of Executives in 1981: 12,023					
Reporting Level	Firm Tenure	Job Tenure	Age (median)	Year of Schooling	
all	**15.2**	**4.2**	**48.0**	**16.4**	
1	22.7	6.8	57.0	17.0	
2	16.0	4.3	51.0	17.0	
3	14.6	4.1	49.0	16.8	
4	14.6	4.1	47.0	16.4	
5	15.1	4.0	47.0	16.2	
6	16.0	4.1	6.5	15.8	
7	18.2	4.5	47.0	15.2	
8	17.1	4.3	6.5	15.0	
9	20.0	3.8	49.0	15.2	
10	16.3	3.0	39.0	13.1	
11	22.8	4.0	50.0	14.0	
B. Compensation Measures by Sampling Year: 1981-85					
Year	1981	1982	1983	1984	1985
No. of Exe	12,023	15,749	18,578	15,611	12,190
Total Pay	$103,962	$105,680	$104,333	$110,854	$120,743
Base	$82,016	$82,982	$85,218	$88,314	$92,626
Bonus	$21,945	$22,698	$19,110	$22,540	$28,118
C. Changes in Firm Characteristics and Job Transition Status by Sampling Year					
Year		1981-1982	1982-1983	1983-1984	1984-1985
Δ%sales		12.6%	3.9%	3.0%	18.5%
Δ%profits		15.2%	24.8%	-12.3%	27.6%
Δ%sizes		2.1%	-0.3%	0.9%	7.8%
Promotion		12.4%	13.4%	14.4%	13.8%
Lateral move		10.6%	12.2%	12.1%	23.5%
Demotion		2.4%	2.1%	2.3%	3.4%

Note: Total pay, base pay and bonus pay are in 1982 US dollars

years. This results in a sample of 290 firms, 19,149 executives, and 74,153 executive-year observations.

Table 1.2 provides summary statistics of the sample to be used in the empirical tests. Panel A shows the individual characteristics of the 1981 cohort. The median age of the executives is 48, which is higher than the median age of the general working population in the U.S.[23] The executives have, on average, 4.2 years of job tenure on the current position and 15.2 years of firm tenure since the first hire. The average years-of-education is 16.4. Most of the executives are in reporting levels 1 through 8. The average firm tenure for a CEO position (reporting level 1) is 23 years. Panel B summarizes different compensation measures for executives from 1981 to 1985. The total number of executives in the sample varies across years. Their average annual real total earnings grew from $103,962 in 1981 (which is equal to $251,560.03 in the 2013 dollar) to $120,743 (or $292,165.53 in the 2013 dollar) in 1985. Bonuses count for approximately 25% of their total pay in each year. From Panel C, in each year, about 12% -14% of the executives are promoted, 10% -12% are laterally moved, and 2.0% -2.8% are demoted.[24] We can see that in this dataset lateral moves are as prevalent as promotions. In addition, 42% of the executives have no lateral moves or promotions across the sample years. About 3.3% of them have more than one promotion within the five years of the sample; while about 7.7% of them have more than one lateral move. In addition, among executives who have a promotion following a lateral move within the sample years, a promotion is earned 1.5 years, on average, after a lateral move.

1.4.2 Empirical Tests

In this subsection, I test the theoretical predictions regarding lateral moves, promotions, wage changes and education. I first present the empirical evidence regarding lateral moves and promotion probabilities (Corollary 1). Then I discuss lateral moves and wage

[23]The median age of the workforce is 42.1 in 2011 according to the Bureau of Labor Statistics. The median working age should be lower in the 1980s given an aging workforce in the US.

[24]The lateral move rate is 23.5% from 1984 to 1985. This spike is caused by an increasing number of unit level changes from 1984 to 1985 and the fact that lateral moves are defined as a within level movement with a unit level or job title change. In the following analysis, only the effect of lateral moves in 1982 and 1983 are considered. So the spike in the 1985 lateral move rate is not a concern.

changes (Corollary 2). The relationship between education and lateral moves (Corollary 3) is investigated last.

Lateral moves and future promotion probabilities

Corollary 1 states that lateral moves in the current period are associated with a higher probability of promotion in the next period. As I discussed before, this positive relationship contains two effects. First, there is a positive selection on workers' expected ability, meaning, higher ability workers (relative to non-movers) are laterally moved. Second, lateral moves have a permanent effect on workers' human capital development and enable them to reach a higher human capital level upon promotion. Thus, lateral moves make the worker more likely to be promoted. The second effect can be understood as the treatment effect of lateral moves.

Empirically, the selection effect should be accounted for by controlling for workers' expected ability. To do this, I use individuals' past period's total pay as a proxy for workers' expected ability. From the theoretical model, firms form expectations about workers' abilities upon observing their outputs. Given a competitive market, workers' outputs are equal to their total pay. Therefore, individuals' past period's total pay is a proxy for expected output (and expected ability). Furthermore, I control for individuals' outputs at the initial level when they first enter the sample.[25] In addition, the theoretical model considers workers who enter the job market (i.e., the lower level jobs) at the same time. This translates into controlling for age and tenure empirically. I also control

[25]Previous studies have used subjective performance evaluation measures as a control for expected output (DeVaro and Waldman, 2012). In more recent studies, Gittings (2012) and DeVaro (2012) have used bonus as a measure of expected output. My data does not contain information on subjective performance. In addition, it is a well-established stylized fact that past wages and wage changes predict promotion (Baker et. al., 1994b; Gibbs and Hendricks, 2004). Thus, I use past period's total compensation as a measure of expected output. I also use other compensation variables as measures of expected output and the results are robust. See Table A.2 in the Appendix for details.

Table 1.3: Lateral Moves and Subsequent Promotions by Year

		Year of Promotion					
		1983		1984		1985	
Year of Lateral Move		No	Yes	No	Yes	No	Yes
		one year after		*two years after*		*three years after*	
1982	Mover	1,078	153	657	132	411	81
		87.6%	12.4%	83.3%	16.7%	83.5%	16.5%
	Non-	7,830	1147	5050	906	3286	462
	mover	87.2%	12.8%	84.8%	15.2%	87.7%	12.3%
				one year after		*two years after*	
1983	Mover			1,011	192	547	108
				84.0%	16.0%	83.5%	16.5%
	Non-			6,918	1,238	6,785	978
	mover			84.8%	15.2%	87.4%	12.6%
						one year after	
1984	Mover					825	117
						87.6%	12.4%
	Non-					5960	861
	mover					87.4%	12.6%

Note: Sample restricted to executives who are not promoted in the current period

for individuals' years of schooling and initial functional area to capture the general education level and the education type in the theoretical model.

In practice, it is possible that promotion is not immediate after an initial move since it may take several periods to develop the task-specific human capital for the upper-level job and different individuals need different amounts of time to develop human capital. Therefore, a more realistic interpretation of Corollary 1 is that a lateral move is positively correlated with promotion probabilities in periods after the move (not necessarily the first period after the move).

Table 1.3 looks at lateral moves and their impact on future promotions in different years. After the initial move, 12.4% of the 1982-lateral-movers are promoted in 1983 while a slightly higher proportion (12.8%) of the executives who stayed in the same position in 1982 are promoted in 1983. However, two years after the lateral move in 1982, among those who were not previously promoted, 16.7% of the 1982-lateral-movers are promoted in 1984 while only 15.2% of the 1982-non-movers are promoted. This difference is even larger in 1985 - three years after the lateral move in 1982, 16.5% of the movers are promoted while only 12.3% of the non-movers are promoted. This pattern repeats for lateral moves in 1983.

To explore the effect of lateral moves on subsequent promotions, I estimate a reduced form model of promotion in which the probability of promotion is a function of individual and firm characteristics and is affected by a previous lateral move. Rather than pooling all lateral moves together, I consider lateral moves in different years separately.[26] The probability that executive i in firm m is promoted in period t is defined by the following equation:

$$Prob(Promotion_{imt} = 1) = \\ F(\delta Lateral_{im\tau} + \beta_r rtotal_{imt-1} + \beta_p Level_{imt-1} + \beta_F F_{mt-1} + \beta_U Unemp_{t-1} + C_{im}). \tag{1.4}$$

$F(\cdot)$ is a cumulative distribution function. $Lateral_{im\tau}$ is a dummy variable that is equal to one if an executive has been laterally moved in period $\tau(\tau < t)$ and zero otherwise. $rtotal_{imt-1}$ is the total compensation in the previous period, which is the measure of expected ability. $Level_{imt-1}$ is a set of dummies indicating the executives reporting level in $t-1$ (I exclude CEO (level 1) positions in the lateral-move year since CEO is the highest position for promotion). F_{mt-1} is a set of firm-specific variables, including promotion opportunities, profits, sales and firm's total employment in period $t-1$. Promotion opportunities are defined as the percentage of executives hired from outside into positions above a given individual. I also include changes of those firm characteristics

[26] If a pooled regression is considered, I still need to control for lateral moves in different years and the time between lateral moves and promotions. This gives the same subsamples as I estimate separately. In addition, since there is no the third-year after the move for lateral moves in 1983 (because the sample ends in 1985), all the third-year effect of lateral moves on promotions comes from lateral moves in 1982. Thus, by partitioning the sample by lateral moves in different years, I can present the result more clearly.

from period $t-1$ to period t. $Unemp_{t-1}$ is the unemployment rate in period $t-1$, which captures the overall labor market conditions. C_{im} captures other individual characteristics. It consists of an observable component and an orthogonal unobserved component. The specification takes the following form:

$$C_{im} = c_X X_{im0} + c_r rtotal_{im0} + c_U Unemp_{imh} + c_A FA_{im0} + u_{im}. \tag{1.5}$$

X_{im0} is a set of human capital measures, including age, education, job tenure, and firm tenure all measured at the point of time when individuals first enter into the sample. $rtotal_{im0}$ is the real total pay when an executive first enters into the sample. This measure controls for the pre-in-sample individual heterogeneity in expected productivity. $Unemp_{imh}$ is the unemployment rate when an executive is hired, which captures potential cohort effects upon hiring that might affect executives' later career development. FA_{im0} represents executives' initial functional area at the time when they are observed. It captures the initial matching effect (α_{ij}) as specified in the theoretical analysis. u_{im} denotes the orthogonal unobserved component.

Conditional on individuals' expected ability, the "treatment effect" of lateral moves on promotions is captured by the lateral move dummy. Table 1.4 presents results from estimating equation (1.4) using a Linear Probability Model (LPM). Column (1) fits a pooled-OLS model. As we can see, a lateral move in 1982 increases the average probability of promotion in years after the move by 1.7%. To further control for unobserved individual heterogeneity, in column (2), I fit a LPM with Random Effect. The effect of lateral moves is still positive but not statistically significant.

While lateral moves exhibit positive effects on future promotions in the full sample, the results might be biased since the 1982 and 1983 entry cohorts are not "at risk" to

Table 1.4: Lateral Moves and Promotion Probability After the Moves: Linear Probility Model

	Lateral Moves in 82 (τ=82; t=83,84,85)				Lateral Moves in 83 (τ=83; t=84,85)			
Executives' Entry Year	*81,82,83 (full-sample)*		*81(restricted-sample)*		*81,82,83 (full-sample)*		*81,82 (restricted-sample)*	
Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Promotion (t)=1 if Yes	OLS	RE	RE	RE	OLS	RE	RE	RE
Lateral (τ)	0.017**	0.014	0.007	-	0.013*	0.008	0.017	-
	(0.007)	(0.010)	(0.010)	-	(0.008)	(0.010)	(0.010)	-
Lateral move effect by year								
Lateral (τ)*Year 1	-	-	-	-0.041***	-	-	-	-0.004
	-	-	-	(0.011)	-	-	-	(0.011)
Lateral (τ)*Year 2	-	-	-	0.053***	-	-	-	0.064***
	-	-	-	(0.013)	-	-	-	(0.014)
Latearl (τ)*Year 3	-	-	-	0.092***	-	-	-	-
	-	-	-	(0.015)	-	-	-	-
Past in-sample observables								
Total Pay (10k)(t-1)	0.004***	0.007***	0.004***	0.004***	0.005***	0.009***	0.008***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Levels (t-1)								
Level 2	-0.428***	-0.568***	-0.486***	-0.481***	-0.412***	-0.513***	-0.472***	-0.469***
	(0.013)	(0.017)	(0.017)	(0.017)	(0.016)	(0.020)	(0.020)	(0.020)
Level 3	-0.314***	-0.427***	-0.351***	-0.346***	-0.309***	-0.389***	-0.340***	-0.337***
	(0.011)	(0.013)	(0.012)	(0.012)	(0.013)	(0.016)	(0.014)	(0.014)
Level 4	-0.227***	-0.315***	-0.244***	-0.240***	-0.210***	-0.270***	-0.219***	-0.217***
	(0.010)	(0.012)	(0.011)	(0.011)	(0.012)	(0.014)	(0.012)	(0.012)
Level 5	-0.131***	-0.193***	-0.134***	-0.132***	-0.114***	-0.154***	-0.102***	-0.102***
	(0.010)	(0.012)	(0.011)	(0.011)	(0.012)	(0.014)	(0.012)	(0.012)
Level 6	-0.036***	-0.067***	*Ref.*	*Ref.*	-0.017	-0.034**	*Ref.*	*Ref.*
	(0.011)	(0.012)	-	-	(0.013)	(0.015)	-	-
Level 7+	*Ref.*	*Ref.*	-	-	*Ref.*	*Ref.*	-	-
Unemp. Rate (t-1)	0.014	0.017**	0.012	0.012	-0.004	-0.003	-0.002	-0.001
	(0.007)	(0.007)	(0.009)	(0.009)	(0.008)	(0.009)	(0.011)	(0.011)

(Table continues)

(Table continues from the previous page)

Firm Char. (t-1)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Individual Initial Char.								
Education	0.007***	0.010***	0.008***	0.008***	0.008***	0.010***	0.009***	0.009***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Age	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Job Tenure	-0.002***	-0.002***	-0.002**	-0.002**	-0.003***	-0.004***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Employer Tenure	0.000	0.000	-0.000	-0.000	-0.000	-0.000	-0.001*	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unemp Rate at hiring	-0.015**	-0.015**	-0.012	-0.012	0.002	0.001	-0.001	-0.002
	(0.007)	(0.007)	(0.009)	(0.009)	(0.008)	(0.009)	(0.011)	(0.011)
Total Pay(10k)(initial)	0.003***	0.003***	0.004***	0.005***	0.002***	0.000	0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Functional Areas	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Constant	0.271***	0.328***	0.326***	0.326***	0.294***	0.347***	0.358***	0.358***
	(0.026)	(0.035)	(0.042)	(0.042)	(0.032)	(0.041)	(0.047)	(0.047)
σ_u	-	0.258	0.249	0.248	-	0.278	0.267	0.266
σ_e	-	0.269	0.265	0.263	-	0.245	0.237	0.236
ρ	-	0.479	0.469	0.471	-	0.564	0.560	0.560
N. of Observations	29,503	29,503	19,261	19,261	19,765	19,765	13,451	13,451
N. of firm-individual	-	14,097	9,586	9,586	-	12,144	8,644	8,644

Note 1 - Standard errors are in parentheses.
Note 2- Sample restricted to executives who are not promoted in the current period.
* Statistically significant at the 10% level.
** Statistically significant at the 5% level.
*** Statistically significant at the 1% level.

lateral moves in 1982. For example, when we consider the effect of lateral moves in 1982 on 1983's promotions, only individuals who enter the sample in 1981 are subject to a lateral move in 1982. Individuals who enter the sample in 1982 and 1983 are automatically counted as non-lateral-movers. These entrants may increase the probability that a non-mover is not promoted (if those new entrants in 1982 are not promoted in 1983) and thus exaggerate the effect of lateral moves on promotions. To make sure we measure the effect of lateral moves on individuals who are subject to lateral moves, in columns (3) and (4) I restrict the sample to individuals who enter the sample in 1981 such that they are subject to lateral moves in 1982. Similarly, in column (7) and (8) I restrict the observations to individuals who enter the sample in either 1981 or 1982 such that they are subject to lateral moves in 1983. I further restrict the sample to levels 2 to 6 in the period when the lateral moves occur.[27]

In the restricted sample, lateral moves continue to exhibit positive impact on future promotions but the magnitude is much smaller than that in the full sample for the 1982 moves. For example, column (3) tells us that a laterally moved worker in 1982 is 0.7% more likely to earn a promotion in periods after 1982. However, this effect is not statistically significant. As I argued in the previous section, the effect of lateral moves on promotions might not be immediate after the move. Therefore, in column (4) I allow the effect of lateral moves to vary over time. To be specific, I include interaction terms of lateral move status by year-after-the-move. We can see that lateral moves exhibit strong positive correlation with promotion probabilities two years after the move. In particular, a lateral move in 1982 increases the probability of promotion in 1984 - two years after the move - by 5.3%. It increases the probability of promotion in 1985 - three years after the move - by 9.2%. However, there is a negative effect of lateral moves on promotions one year after the move. One explanation is that it may take several years for a moved

[27]Table A.1 summarizes lateral moves in each year by level. We can see that there is a sudden drop in the number of lateral moves beyond level 6.

Table 1.5: Lateral Moves and Promotion Probability After the Moves: Probit Model

	Lateral Move in 82 (τ=82;t=83,84,85)			Lateral Moves in 83 (τ=83;t=84,85)		
Exec's Entry Year	*81(restricted sample)*			*81,82(restricted sample)*		
Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)
Promotion (t)=1 if Yes	Probit	Probit-RE	Probit-RE	Probit	Probit-RE	Probit-RE
Lateral (τ)	0.009	0.009	-	0.016*	0.016*	-
	(0.007)	(0.007)	-	(0.009)	(0.009)	-
Lateral move effect by year						
Lateral (τ)*Year1	-	-	-0.016*	-	-	0.009
	-	-	(0.009)	-	-	(0.010)
Lateral (τ)*Year2	-	-	0.034***	-	-	0.031**
	-	-	(0.013)	-	-	(0.015)
Latearl (τ)*Year3	-	-	0.032*	-	-	-
	-	-	(0.017)	-	-	-
Past in-sample observables						
Total Pay (10k)(t-1)	0.004***	0.004***	0.004***	0.007***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Levels (t-1)						
Level 2	-0.334***	-0.356***	-0.355***	-0.332***	-0.354***	-0.354***
	(0.012)	(0.012)	(0.012)	(0.013)	(0.015)	(0.015)
Level 3	-0.284***	-0.307***	-0.307***	-0.291***	-0.322***	-0.323***
	(0.012)	(0.012)	(0.012)	(0.013)	(0.016)	(0.016)
Level 4	-0.204***	-0.223***	-0.223***	-0.199***	-0.230***	-0.233***
	(0.012)	(0.012)	(0.012)	(0.014)	(0.017)	(0.017)
Level 5	-0.107***	-0.118***	-0.117***	-0.090***	-0.107***	-0.110***
	(0.013)	(0.013)	(0.013)	(0.014)	(0.017)	(0.017)
Level 6	*Ref.*	*Ref.*	*Ref.*	*Ref.*	*Ref.*	*Ref.*
Unemp. Rate (t-1)	0.009	0.009	0.009	-0.001	-0.002	-0.002
	(0.008)	(0.008)	(0.008)	(0.011)	(0.011)	(0.011)

(Table continues)

(Table continues from the previous page)

Firm Char. (t-1)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Individual Initial Char.						
Education	0.005***	0.005***	0.005***	0.007***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Age	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Job Tenure	-0.002***	-0.002***	-0.002***	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Employer Tenure	-0.000	-0.000	-0.000	-0.001**	-0.001**	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unemp Rate at hiring	-0.010	-0.010	-0.010	-0.002	-0.002	-0.002
	(0.008)	(0.008)	(0.008)	(0.011)	(0.011)	(0.011)
Total Pay(10k)(initial)	0.004***	0.004***	0.004***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Functional Areas	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
σ_u	-	0.00165	0.00379	-	0.423	0.480
ρ	-	2.71e-06	1.44e-05	-	0.152	0.187
N. of Observations	19,261	19,261	19,261	13,451	13,451	13,451
N. of firm-individuals	-	9,586	9,586	-	8,644	8,644
N. of quadrature points	-	12	12	-	12	12
Log likelihood	-6596	-6596	-6589	-4603	-4602	-4601

Note 1 - Standard errors are in parentheses. Average Marginal Effects are reported. Marginal effects for dummies are the differences in predicted probabilities when the dummy equals 1 and when it equals 0.
Note 2 - Sample restricted to executives who are not promoted in the current period.
* Statistically significant at the 10% level.
** Statistically significant at the 5% level.
*** Statistically significant at the 1% level.

individual to develop sufficient human capital in the new position to warrant promotion. Given that the average length of a lateral-move-promotion spell in this sample is 1.5 years, it is reasonable that we do not see a positive effect of lateral moves on promotions one year after the move. The same pattern repeats with the 1983-lateral-moves. The results in column (4) and (8) suggest that we need to allow the effect of the move to differ over time because the average result across years does not fully capture the effect of lateral moves on promotion.[28]

The LPM provides some support for the prediction that lateral moves are positively associated with promotions. In Table 1.5, I estimate equation (1.4) with a non-linear model (applied to the restricted sample only). To keep the estimation strategy straightforward, I assume the orthogonal component in the individual heterogeneity (i.e. u_{im} in equation (1.5)) follows a normal distribution and thus I implement a Random Effect Probit model using Butler and Moffitt's (1982) method. I fit a pooled Probit model in columns (1) and (5) and a Probit model with Random Effect in columns (2) and (6) controlling for individual characteristics. The effect of lateral moves in either 1982 or 1983 is positive but insignificant.

In columns (3) and (6), I allow the effect of lateral moves to vary over time. The results are consistent with those in the LPM. Lateral moves are associated with a 3.1% - 3.4% increase in the marginal probability of promotion two years after the move. There is a negative effect associated with the 1982-moves one year after the move. Note that the effects of lateral moves might be underestimated since we do not observe the full employment history. For example, when we consider the effect of a 1982-lateral-move on 1983-promotion-probability, a promoted worker in 1983 who is not moved in 1982 but is moved in 1981 will add to the probability of observing a non-mover in 1982 being

[28]The results from pooled OLS Models (instead of the Random Effect Models) are very similar to those in Table 1.4, columns (3), (4), (7) and (8). The estimates are available up on request.

promoted in 1983, but in fact the 1983-promotion is due to an unobserved 1981-move. Since I do not observe the whole employment history of a worker after they enter the firm, I cannot eliminate this bias.[29]

Overall, lateral moves are positively correlated with subsequent promotions. The positive effect is strong two years after the move. This result suggests that it may take an executive more than one year after a lateral move to acquire enough task-skills to warrant subsequent promotions.[30]

Lateral moves and future wage growth

In this subsection, I investigate the relationship between lateral moves and wage changes (Corollary 2). Figure 1.1 plots the raw compensation data against years by lateral-move status in 1982 and 1983. The top two graphs compare the wage profiles of the 1982-lateral-movers and the 1982-non-movers. Without controlling for other covariates, the average total pay and the base pay of the 1982-lateral-movers are below those of the 1982-non-movers. There is no obvious difference in the rates of wage growth between the 1982-movers and the 1982-non-movers from 1981 to 1982. However, lateral movers' wages start to grow faster in years after the move. The bottom two graphs plot the total pay and the base pay of the 1983-movers and the 1983-non-movers. In 1983, there is not much difference in terms of the total pay and the base pay between the movers and the non-movers, but the movers' wages grow much faster than the non-movers' after 1983. These raw plots are consistent with Corollary 2 that wage growth for movers is larger in the post-move periods.

[29] I also cluster standard errors in the non-linear model at both individual and firm levels. The results are consistent with what I find in Tables 1.4 and 1.5. See Table A.3 in the Appendix for details.

[30] This is a discrepancy between the current version of the model and the empirical findings. If I extend the model to more than three periods, I can capture this specific timing. However, all other results concerning lateral moves and promotions should stay the same.

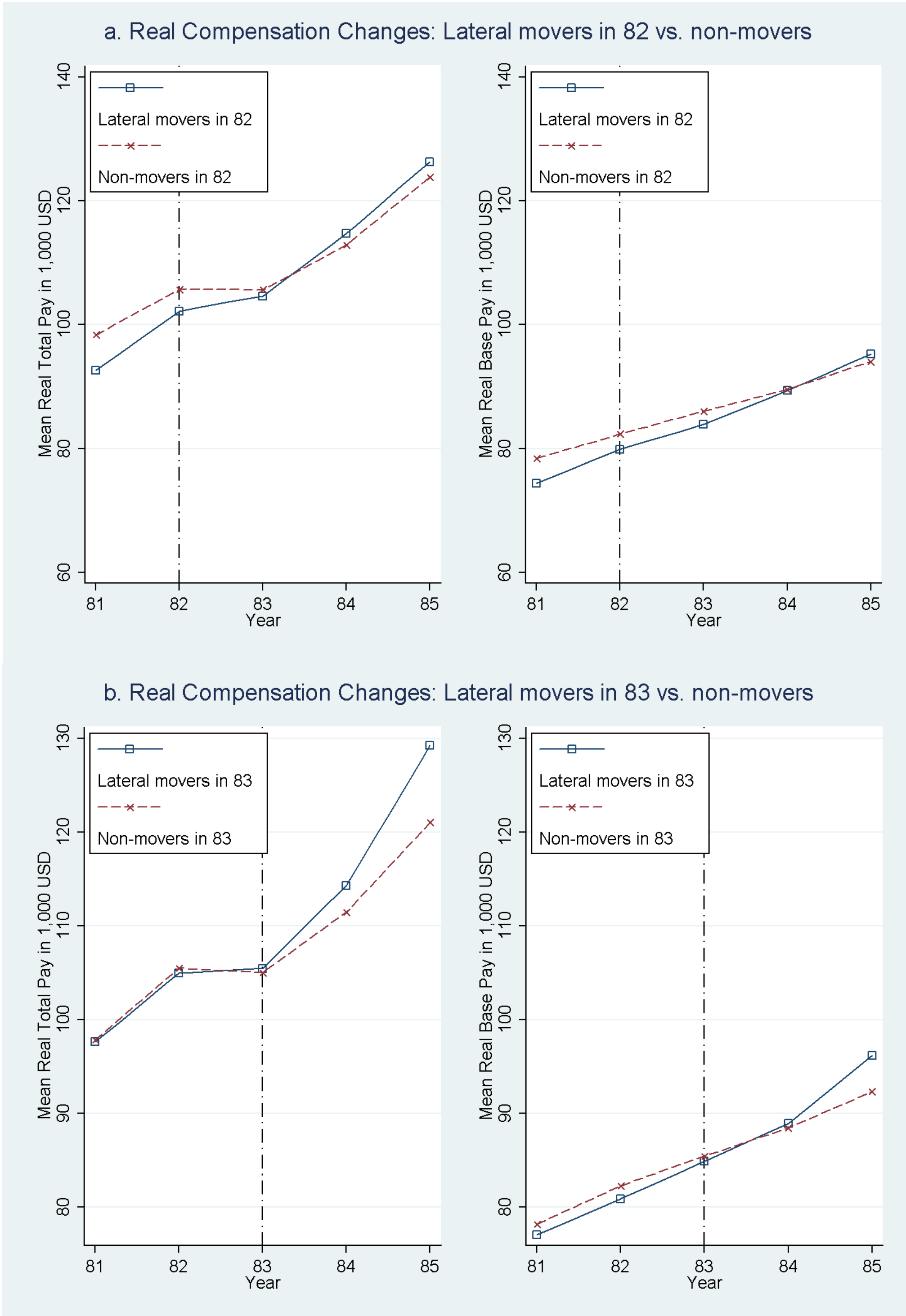

Figure 1.1: Real Compensation changes and lateral move status in 1982 and 1983.

Table 1.6: Lateral Moves and Real Compensation Changes (in 1982 dollar) after the Moves: Pooled OLS

	Lateral Moves in 82 (τ=82; t=83,84,85)			Lateral Moves in 83 (τ=83; t=84,85)		
Real Comp. Change from the Year of the Moves:	(1) Total (t-τ)	(2) Base (t-τ)	(3) Bonus (t-τ)	(4) Total (t-τ)	(5) Base (t-τ)	(6) Bonus (t-τ)
Lateral move effect by year						
Lateral (τ)*Year1	-2,502.459***	-2,056.176***	-446.284	335.122	-980.802***	1,315.923**
	(661.544)	(247.700)	(573.534)	(673.116)	(316.055)	(551.503)
Lateral (τ)*Year2	3,837.632***	1,730.142***	2,107.490***	6,021.859***	2,622.199***	3,399.660***
	(953.559)	(412.555)	(790.310)	(1,276.022)	(505.214)	(1,040.359)
Latearl (τ)*Year3	8,474.815***	4,626.428***	3,848.386***	-	-	-
	(1,703.592)	(741.077)	(1,272.435)	-	-	-
Levels (t)						
Level 1	64,291.996***	57,915.917***	6,376.082	54,839.843**	53,103.655***	1,736.186
	(15,012.208)	(7,607.917)	(11,006.935)	(22,086.481)	(8,465.347)	(19,754.505)
Level 2	10,426.695***	8,137.218***	2,289.477*	12,861.563***	6,724.133***	6,137.430***
	(1,706.119)	(817.994)	(1,280.828)	(1,646.418)	(667.133)	(1,368.044)
Level 3	3,450.460***	2,902.715***	547.745	5,300.007***	2,214.054***	3,085.953***
	(782.028)	(391.294)	(564.558)	(672.115)	(266.680)	(583.375)
Level 4	1,300.450**	1,115.022***	185.428	3,002.916***	835.032***	2,167.884***
	(571.752)	(309.697)	(383.077)	(420.130)	(178.972)	(369.895)
Level 5	123.867	438.984	-315.118	994.166***	182.779	811.387***
	(498.190)	(286.596)	(316.101)	(338.556)	(150.478)	(283.970)
Level 6+	*Ref.*	*Ref.*	*Ref.*	*Ref.*	*Ref.*	*Ref.*

(Table continues)

(Table continues from the previous page)

Δ Firm Char. (t)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Δ Unemp. Rate (t)	-3,834.026**	-1,785.783*	-2,048.244***	-1,419.205	-4.683	-1,414.522*
	(1,656.140)	(1,073.113)	(737.909)	(1,456.334)	(841.954)	(831.854)
Individual Initial Char.						
Education	662.852***	310.176***	352.676***	692.459***	95.369	597.090***
	(153.075)	(70.628)	(122.883)	(150.562)	(63.788)	(123.954)
Age	-48.118	272.531**	-320.649	376.652	33.044	343.607
	(288.446)	(129.620)	(235.109)	(258.250)	(107.021)	(214.077)
Age^2	-1.442	-4.766***	3.325	-6.337**	-1.935	-4.402*
	(3.127)	(1.393)	(2.579)	(2.910)	(1.187)	(2.428)
Job tenure	-418.610***	-102.993***	-315.618***	-324.451***	-140.216***	-184.234***
	(72.640)	(33.129)	(61.845)	(73.171)	(30.100)	(59.821)
Emp. tenure	101.401***	5.264	96.137***	112.343***	35.909***	76.434**
	(35.777)	(15.328)	(29.058)	(35.067)	(12.934)	(29.764)
Total Pay (10k) (initial)	-54.720	321.543***	-376.263***	735.116***	326.326***	408.790***
	(144.413)	(52.777)	(119.173)	(123.495)	(38.370)	(112.262)
Functional Areas	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Constant	-1,948.069	-6,153.693*	4,205.623	-15,398.538**	1,640.460	-17,038.995***
	(7,073.403)	(3,233.428)	(5,690.923)	(6,402.642)	(2,665.871)	(5,309.873)
Observations	19,416	19,416	19,416	13,512	13,512	13,512
R-squared	0.068	0.172	0.046	0.170	0.197	0.105

Note - Robust standard errors are in parentheses; clusted at firm-individual level
Sample restricted to executives who are not promoted in the current period.
* Statistically significant at the 10% level.
** Statistically significant at the 5% level.
*** Statistically significant at the 1% level.

Since the effect of lateral moves on wage growth is realized upon promotion and the effect of lateral moves on promotion probabilities might not be immediate, we would expect the effect of lateral moves on wages to differ over time in practice. Therefore, I allow for different effects of lateral moves on wage changes in different periods after a lateral move.[31] Since wage changes are considered, I only control for changes in firms' characteristics rather than base levels. I estimate the wage change by the following specification:

$$\begin{aligned} & w_{imt} - w_{im\tau} \\ & \quad = \eta Lateral_{im\tau} + \psi_p Level_{imt} + \psi_F \Delta F_{mt} + \psi_U \Delta Unemp_t + \tilde{C}_{im}. \end{aligned} \tag{1.6}$$

$Lateral_{im\tau}$ is a vector of dummies indicating lateral move status by year-after-the-move. For lateral moves in 1982, this vector includes interaction terms of the lateral-move dummy in 1982 with year dummies for 1983, 1984, and 1985. For lateral moves in 1983, this vector includes interaction terms of the lateral-move dummy in 1983 with year dummies for 1984 and 1985. The individual characteristic, $\tilde{C}_{im}$, is defined similarly as in equation (1.4). The only difference is that I include an age-squared term to capture the non-monotonic wage-age profile. Since the theory does not distinguish between base pay and total pay, I estimate equation (1.6) using OLS with clustered standard errors (on firm-individual level) for changes in total compensation, base pay, as well as bonus pay. I consider wage levels rather than log wages because the theoretical model only generates results concerning wage levels.[32]

The results from estimating the wage equation in (1.6) are shown in Table 1.6. Column (1) says that the total compensation change of a 1982-mover from 1982 to 1984 is

[31] Note that the wage difference is calculated as the wage growth from the period of the lateral moves to the current period. This is to match the theoretical model. In the theoretical model, lateral moves affect the expected wage growth from the period of lateral moves to the period of promotion.

[32] Note that in the wage equation, I do not include a variable to control for the expected productivity because by taking the difference in wages, the expected ability cancels out.

$3,838 larger than that of a 1982-non-mover. Similarly, the total compensation change of a 1982-mover from 1982 to 1985 is $8,475 larger than that of a 1982-non-mover. However, a mover has a smaller total compensation change than a non-mover from 1982 to 1983. From Tables 1.4 and 1.5 , we know that a mover is less likely to be promoted than a non-mover one year after the move (recall that there is a negative effect of lateral moves on promotions one year after the moves). Thus, the smaller wage change is due to non-promotion.[33] In contrast, the movers have a larger wage change two years after the move when they are more likely to be promoted relative to the movers. This pattern repeats for different measures of compensation and for lateral moves in 1983.[34]

As discussed in the previous section, while lateral moves are associated with current task-specific human capital losses, it is also associated with positive learning about individuals' innate ability. The net effect of lateral moves on wage growth in the period of lateral move depends on which effect dominates. I estimate a specification similar to equation (1.6) except that the dependent variable is the wage growth in the period of lateral move from the period before the move. The equation I estimate takes the following form:

$$\begin{aligned} & w_{im\tau} - w_{im\tau-1} \\ & \quad = \kappa Lateral_{im\tau} + \gamma_p Level_{im\tau} + \gamma_F \Delta F_{m\tau} + \gamma_U \Delta Unemp_\tau + \tilde{C}_{im} + \varepsilon_{im\tau}. \end{aligned} \tag{1.7}$$

Table 1.7 shows that lateral movers in either 1982 or 1983 have larger pay increases in the period of lateral moves despite that we lose some significance in the bonus measure. The positive relationship between lateral moves and wage growth in the period of lateral moves suggests that the positive learning associated with a lateral move domi-

[33]Since the theory does not distinguish between promoted movers and non-promoted movers, I do not control for promotion status in the wage equation.

[34]As robustness checks, I also cluster the standardard error at the firm level and estimate a Random Effect model. The results are consistant with the estimates in Table 1.6. See Table A.4 in the Appendix for details.

Table 1.7: Lateral Moves and Real Compensation Changes Upon Moving:OLS

	Lateral Moves in 82 (τ=82)			Lateral Moves in 83 (τ=83)		
	(1) ΔTotal(τ)	(2) ΔBase(τ)	(3) ΔBonus(τ)	(4) ΔTotal(τ)	(5) ΔBase(τ)	(6) ΔBonus(τ)
Lateral (τ)	2,506.603***	1,917.978***	588.625	1,931.442**	1,410.541***	520.902
	(730.801)	(291.309)	(642.485)	(787.682)	(390.864)	(588.377)
Levels (τ-1)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Δ Firm Char. (τ)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Δ Unemp. Rate (τ)	10,104.750	7,015.898	3,088.852	-9,866.078**	-5,476.332*	-4,389.746**
	(7,283.072)	(4,678.745)	(2,648.075)	(4,862.081)	(2,931.993)	(2,042.549)
Individual Initial Char.						
Education	348.916**	59.503	289.412**	349.788**	262.283***	87.505
	(148.657)	(57.301)	(127.628)	(142.882)	(57.422)	(116.379)
Age	-386.511	-169.746	-216.765	560.803	465.434**	95.369
	(272.987)	(124.591)	(251.450)	(427.062)	(201.372)	(275.457)
Age^2	4.444	1.064	3.380	-6.141	-5.634***	-0.507
	(3.009)	(1.335)	(2.793)	(4.420)	(2.060)	(2.939)
Job tenure	-216.701***	-99.290***	-117.411**	-436.089***	-85.885*	-350.203***
	(65.729)	(26.152)	(54.666)	(95.364)	(45.391)	(64.532)
Emp. tenure	-61.170*	25.145*	-86.315***	-20.454	-7.459	-12.994
	(36.245)	(13.386)	(31.398)	(34.141)	(15.033)	(27.815)
Total pay (10k)	88.624	257.855***	-169.230*	-586.908***	12.040	-598.948***
	(121.557)	(60.921)	(99.947)	(169.016)	(79.302)	(112.688)
Functional Areas	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Constant	8,656.362	5,993.681**	2,662.680	-10,280.161	-11,015.194**	735.032
	(6,599.523)	(2,996.623)	(6,018.779)	(10,266.710)	(4,844.186)	(6,585.610)
Observations	8,640	8,640	8,640	9,169	9,169	9,169
R-squared	0.058	0.122	0.032	0.066	0.064	0.069

Note - Robust standard errors are in parentheses; clusted at firm-individual level
Sample restricted to executives who are not promoted in the current period.
* Statistically significant at the 10% level. ** Statistically significant at the 5% level.
*** Statistically significant at the 1% level.

nates the potential human capital losses.

Education and lateral moves

I now examine the relationship between education and lateral moves. As discussed in Corollary 3, there is a negative relationship between the probability of a lateral move and the education level.[35] To test this, I estimate a model that is very similar to equation (1.4) except that the dependent variable now is a dummy variable that is equal to one if an individual has an in-sample lateral move in year τ and zero otherwise. I include a full set of dummies for years of schooling. The reference group is years of schooling below 16 years. To be specific, for the executives who are not promoted in period τ, the probability that executive i in firm m is laterally moved is defined by the following equation:

$$Prob(Lateral_{im\tau} = 1) = \quad (1.8)$$
$$F(\pi_{e_1} educ_{im\tau} + \pi_r rtotal_{imt-1} + \pi_p Level_{im\tau-1} + \pi_F F_{m\tau-1} + \pi_U Unemp_{\tau-1} + \hat{C}_{im}).$$

$educ_{im\tau}$ is a set of education dummies. The covariates in $\hat{C}_{im}$ are similar to those specified in equation (1.5) except that the education term is not included. Table 1.8 presents the results from estimating equation (1.8) using both a Random Effect LPM and a Random Effect Probit Model. As shown in columns (1) and (2), there is a strong negative relationship between lateral moves and years of schooling after 19 years. This relationship is more significant in the Probit model than in the linear probability model. This may due to the fact that the linear probability model is not a very good approximation for extreme values. There is a positive but statistically insignificant relationship

[35] In the data, no individuals with years of schooling higher than 21 are laterally moved, which matches the theoretical prediction. However, since only a small number of individuals have more than 21 years of schooling in my sample, it is possible that the true probability of a lateral move for an individual with more than 21 years of schooling is positive.

Table 1.8: Education and Lateral Moves

Dependent Variable:	(1)	(2)	
Lateral move (τ)=1 if yes	LPM-RE	Probit-RE	Margins
Years of Education (τ) =			
<16	*Ref.*	*Ref.*	*Ref.*
16	0.004	0.018	0.004
	(0.006)	(0.026)	(0.006)
17	0.007	0.032	0.008
	(0.008)	(0.034)	(0.008)
18	0.008	0.030	0.007
	(0.008)	(0.033)	(0.008)
19	-0.025**	-0.152***	-0.032***
	(0.011)	(0.049)	(0.010)
20	-0.022*	-0.109**	-0.024**
	(0.011)	(0.049)	(0.010)
21	-0.153	-4.859	-0.157***
	(0.152)	(367.046)	(0.006)
22	-0.138	-4.582	-0.157***
	(0.214)	(540.805)	(0.006)
Past in-sample observables			
Total Pay (10k)(τ-1)	0.006***	0.030***	0.007***
	(0.001)	(0.004)	(0.001)
Levels (τ-1)	*Yes*	*Yes*	*Yes*
Unemp. Rate (τ-1)	0.016***	0.074***	0.017***
	(0.005)	(0.021)	(0.005)
Firm Characteristics (τ-1)	*Yes*	*Yes*	*Yes*
Individual Initial Char.			
Age	-0.003***	-0.013***	-0.003***
	(0.000)	(0.001)	(0.000)
Job Tenure	-0.002***	-0.011***	-0.003***
	(0.001)	(0.003)	(0.001)
Employer Tenure	0.000	0.001	0.000
	(0.000)	(0.001)	(0.000)
Unemp Rate at hiring	-0.017***	-0.076***	-0.017***
	(0.005)	(0.021)	(0.005)
Total Pay/Age (initial)	-0.006***	-0.026***	-0.006***
	(0.001)	(0.004)	(0.001)
Functional Areas	*Yes*	*Yes*	*Yes*
Constant	0.357***	-0.238***	-
	(0.017)	(0.074)	-
σ_u	0.116	0.411	-
σ_e	0.354	-	-
ρ	0.0972	0.145	-
N. of Observations	45,212	45,212	
N. of firm-individuals	18,139	18,139	
N. of quadrature points	-	12	
Log likelihood	-	-19883	

Note - Standard errors are in parentheses; clusted at firm-individual level
Sample restricted to executives who are not promoted in the current period.
* Statistically significant at the 10% level.
** Statistically significant at the 5% level.
*** Statistically significant at the 1% level.

between lateral moves and years of schooling for schooling years before 19. These patterns are consistent with the model that for workers with very high levels of education, the probability of lateral move is negatively correlated with years of education. For workers with low or median levels of education, the relationship between education and lateral moves is inconclusive and depends on how different the upper level and the lower level jobs are.

1.5 Alternative Explanations

In this section, I discuss two potential alternative explanations for the empirical findings in Section 1.4, (i) contracted job rotation and (ii) worker-job matching through lateral moves.

Job rotation is a special type of lateral move with a pre-determined career path. When job rotation is merit-based, i.e., promising workers are selected into the job rotation program, and is designed to provide employee learning through human capital accumulation, job rotation and the lateral move mechanisms described in my model are in essence the same. Those job rotation programs that improve the match between workers and firms are mainly for entry-level workers. Campion et al. (1994) find that young workers are more likely to engage in job rotations, while upper level managers are more likely to engage in a mechanism that provides "a broader scope of the business operations" to the managers. Since I have a senior sample I argue that matching-type of job rotations are unlikely to be the main contract arrangement for the individuals in my sample.

Another closely related explanation is that lateral moves (without a rotation program) improve worker-job matching. Borrowing from Novos's (1995) idea about how turnover improves worker-job matching, let us consider a modeling environment without human capital accumulation but lateral moves improve worker-job matches. If a worker matches with job A badly, she moves to job B. If the match-quality or job-preference is private information, it is not very intuitive to argue why this worker should be a better candidate for the upper-level job relative to a worker who is a good match for job A. In reality, lateral moves might improve worker-job matching. However, worker-job matching is unlikely to drive the results that we observed in the data such as the positive relationship between promotion probability and lateral moves.

1.6 Conclusion

Lateral moves play an important role in individuals' career development since they facilitate human capital development. However, studies in the literature on career and wage dynamics inside firms focus almost exclusively on promotions. In this essay I extend the theory by incorporating lateral moves into a job assignment model with learning and task-specific human capital accumulation. I also examine how the publicly observed education levels affect individuals' lateral moves. This model generates a set of new testable predictions concerning the relationship between lateral moves and promotion probability, wage changes, as well as education.

I test the model's predictions using a large employer-employee linked panel in the United States. My empirical results support the theoretical model to a large extent. I find that, controlling for tenure and unobserved heterogeneity, laterally moved workers are more likely to be promoted two years after the move compared to workers who stayed in the same position. I also find that laterally moved workers experience a larger wage growth in periods after the move compared to workers who stayed in the same position. In addition, the probability of a lateral move among those who are not promoted decreases with years of education for those with more than 19 years of schooling.

This study can be extended in several directions. Theoretically, I plan on incorporating turnover into the model to consider lateral mobility across firms. Second, I can introduce some asymmetry in the production process for the lower level jobs to analyze different promotion paths (Clemens, 2012; Sasaki et al., 2012; Bezil et al., 2012). Empirically, I want to test the relationship between MBA training and lateral moves. Murphy and Zabojnik (2006) show that there are an increasing number of newly appointed CEOs with an MBA degree. If MBA training can substitute for task-skills (especially general management skills) acquired on the job, I expect a decrease in the number of lateral moves before promotion to the CEO position in later years.

CHAPTER 2

THE SIGNALING ROLE OF NOT BEING PROMOTED: THEORY AND EVIDENCE

2.1 Introduction

Promotion sends positive signals about a worker's ability and productivity. This theoretical possibility has been extensively studied in a body of literature stemming from Waldman (1984a), who formulates a two-job-level two-period model. Waldman assumes that a worker's current employer has better information about a worker's true ability by observing this worker at work. Potential employers can only infer the worker's ability by observing her current period's job assignment made by the incumbent firm. Waldman's two main conclusions are (1) promotions send positive signals about workers' abilities and thus are associated with substantial wage increases; (2) the incumbent firm promotes fewer employees than what is socially optimal and this distortion is more severe when workers' human capital is general rather than firm specific.

While the signals associated with promotion have been extensively examined, the signals associated with non-promotion, on the other hand, are surprisingly understudied. In this essay, I extend the promotion-as-signal framework by arguing that additional years of job-level tenure (i.e., non-promotion) sends negative signals about a worker's ability. Intuitively, if a worker stays in the same job level for many years while her peers are all promoted, this worker is believed to be less competent (or is less likely to be a productive worker). These negative beliefs eventually translate into small wage increases or even wage decreases.

Although the basic idea is intuitive, formal theoretical models that explore this negative signaling idea are almost non-existent. Bernhardt (1995) is the only previous study that captures the negative signals associated with not being promoted. In that study, Bernhardt argues that there is a negative sorting in promotion such that abler workers are promoted earlier. However, that analysis makes ambiguous predictions regarding the implications of the negative sorting for workers' wage dynamics. In this essay, I build a T-period model with two job levels to explore the negative signaling role of non-promotion on workers' wage dynamics.

My model shows that when a worker stays in the same job level for a long time, her wages first increase then decrease with additional job-level tenure. To see the logic, consider a set up where a worker's productivity is jointly determined by her expected ability and on-the-job human capital accumulation. Firms use non-promotion as signals to infer workers' abilities. If firms keep on receiving negative signals about a worker's ability from non-promotion, this worker is perceived less likely to be a productive worker. The negative signals associated with additional job-tenure eventually cause the non-promoted workers' wages to fall since productivity rises little with additional job-level tenure after the worker spends a long time on the same job. But since human capital accumulates very fast when a worker first starts on a job, the fast human capital accumulation outweighs the downward adjustment (due to non-promotion) in beliefs about a worker's expected ability and therefore the non-promoted workers' wages rise with job-tenure even though there is negative learning about their abilities.

From the above reasoning, if one only considers the learning component in the wage determination process, wages should decrease with additional job-tenure. On the other hand, if one only considers the human capital component, wages should increase then flatten out with additional tenure on the same job. By bringing together the learning argument and human capital theory, I can explain two wage patterns observed in Baker et al. (1994a). First, pre-promotion wages increase then decrease with job-level tenure. Second, the wages in the periods of promotion increase then decrease with job-level tenure on the lower level job.

This essay contributes to the literature in several different ways. First, it fills a gap in the learning literature by capturing the negative signaling role of non-promotion. Second, it contributes to the human capital literature by exploring the relationship between wages and job-level tenure. Third, it provides an explanation for a set of empirical findings that are not well captured in existing models. It also provides empirical evidence that is consistent with my model's predictions.

The outline of the essay is as follows. The next section reviews the related literature. Section 3 sets up the model. In Section 4, I first analyze a T-period model with full information then compare equilibrium behavior in this benchmark model to equilibrium behavior in a model with asymmetric information. I present empirical evidence in Section 5. Section 6 concludes.

2.2 Related Literature

This essay connects two theoretical building blocks in the existing literature on wage and career dynamics inside firms - learning and on-the-job human-capital accumulation.

The learning literature falls into two broad categories. One set of papers assumes symmetric learning where workers' outputs are observed by all firms in the market (Harris and Holmstrom, 1982). The other set of papers assumes asymmetric learning where a worker's current employer privately observes the worker's output. The asymmetric learning literature further divides into two areas of focus. One set of papers investigates the adverse selection issue in labor market turnover following Greenwald (1986). The other set of papers explores the idea of promotions as signals following Waldman (1984a). My essay builds on the promotion-as-signal approach under asymmetric learning.

The promotion-as-signal approach has been extended in many different ways. Ricart-i-Costa (1988) considers a two-period n-job-level model. Bernhardt (1995) considers a two-level model with infinite periods. In Zabojnik and Bernhardt (2001), a promotion signal contains information about workers' human-capital investment rather than their innate ability. DeVaro and Waldman (2012) consider how promotion signals vary with education.[1] The promotion-as-signal approach has also been extended to analyze up-or-out contracts and turnover (Bernhardt and Scoones, 1999; Ghosh and Waldman, 2010). Waldman (1984a) and the various extensions capture many stylized facts about wage and promotion dynamics, such as large wage increases upon promotion (Bernhardt, 1995) and the wage-and-firm-size effect (Zabojnik and Bernhardt, 2001), etc. Yet, these studies have almost exclusively focused on the positive signals associated with promotion. My essay fills a gap in the learning literature by exploring the negative signals associated with non-promotion.

Another important perspective concerning workers' wage and career dynamics inside firms is on-the-job human-capital acquisition. Numerous empirical studies have investigated the contribution of firm tenure and total labor market experience to individuals' wage growth and find a concave wage-firm-tenure profile using household surveys (Altonji and Shakotko, 1987; Topel, 1991; Altonji and Williams, 2005; Sullivan, 2010; Pavan, 2011).[2] On the other hand, using a 1% sample of the British labor force, Dev-

[1] DeVaro and Waldman (2012) treat education as a measure of initial human capital stock, not as another source of signals.

[2] Sanders and Taber (2012) provide a comprehensive survey of the literature on life-cycle wage growth and heterogeneous human-capital accumulation.

ereux et. al. (2013) find a hump-shaped relation between wages and job-tenure when they do not control for firm-tenure.[3] Using firms' personnel records, Baker at al. (1994a) also find a hump-shaped wage-job-tenure profile for non-promoted workers, i.e., their wages first increase with job-tenure then decrease.

While standard human capital accumulation theory can explain the increase of wages with job-tenure when workers are new to a job, it cannot explain why wages fall when workers stay on the same job for a long time. The asymmetric learning framework with human capital accumulation in Waldman (1984a) might potentially explain the hump-shaped wage-job-tenure relation, however, neither the original '84 model nor most of the later extensions in the promotion-as-signal literature capture the negative signals of non-promotion. There are two reasons. First, those models assume that a worker's current employer learns about the worker's ability perfectly after one period of employment. Second, many of those models have a strong "winner's curse" in their equilibrium (Milgrom and Oster, 1987). In models with asymmetric learning, firms only observe the job assignments of the workers at other firms (they still observe the output of their own workers). When the winner's curse occurs, the wage offer a firm is willing to make to a non-promoted worker at another firm is determined by the lowest possible expected ability level among workers with the same job assignment history. Furthermore, this lowest expected ability does not vary with job-level tenure for the non-promoted worker if the current employer learns the worker's ability perfectly after a period. Thus, the signal associated with non-promotion does not cause further adjustments in wage offers to pervious employees and thus wages actually paid.[4] My model moves away from the

[3]They find a negative relation between job tenure and wages holding firm tenure constant. But in their study, they do not control for levels and they do not distinguish between promoted workers and non-promoted workers. So the negative relation between job tenure and wages may capture the wage difference between the promoted workers who have very short job-tenure and the non-promoted workers who in general have longer job-tenure in the current job.

[4]In the '84 model, the output on the lower level job is assumed to be a constant. Thus, we should observe wages stay constant with additional job tenure. In later extensions, workers' outputs are determined by their expected abilities and human capital levels. With a strong winner's curse, we should observe

strong winner's curse problem by assuming the existence of exogenous turnover as in Greenwald (1986). With exogenous movers, outside firms are willing to offer wages that are above the expected productivity of the lowest ability worker within a certain job assignment history. That is, there are further adjustments in wages when a non-promotion is observed.

The only previous paper that captures negative signals associated with non-promotion is Bernhardt (1995). Bernhardt considers a framework with human capital accumulation and asymmetric learning without the winner's curse. However, his model predicts that the non-promoted workers' wages can either increase or decrease with additional tenure but the relationship is monotonic. In addition, Bernhardt focuses on the relationship between wages and firm-tenure. However, the household surveys show that workers' wages do not fall overall with firm-tenure. As shown in Baker et al. (1994a), workers' wages only fall with job-tenure when they stay a long time on the same job level.

In summary, most of the existing promotion-as-signal models do not capture the negative signals of non-promotion. The only theoretical model capturing this idea makes predictions that do not match the evidence. By combining the asymmetric learning argument and human capital theory, I develop a tractable framework to capture a set of empirical findings in Baker et al. (1994a). In particular, Baker et al. show that, for a worker who stays in the same job level for more than six years, her wages first increase then decrease. In addition, if a worker earns a promotion within four or five years of entering into a job level, her wage in the period of promotion is higher than the promotion wage paid to a worker who was promoted in the previous period. But if the promotion is more than four or five years after entering into the job level, her wage upon

wages increase with tenure in those models due to human-capital accumulation because a non-promoted worker's expected output is independent of her ability.

promotion is lower than the promotion wage paid to a worker who was promoted in the previous period.

2.3 The Model

In this section, I set up a T-period model to analyze the role of non-promotion on workers' wage dynamics.

There is free entry into the market. All firms are identical with two job levels. The manager jobs (m) are on the upper level and the laborer jobs (l) are on the lower level .

Workers enter the labor market in period 1. They are either good (g) or ordinary (r). Let A denote workers' ability types, i.e., $A \in \{g, r\}$. Neither the firms nor the workers themselves observe the true type of a particular worker. However, their prior belief is that a worker is good with probability p_0. I assume p_0 is sufficiently small that, given the production function defined below, all workers are assigned to the laborer job in period 1.

Outputs are jointly determined by workers' ability types and their human capital levels. Firms learn about workers' types gradually by observing workers' output realizations. This is a generalization of the setting in previous papers such as Waldman (1984a) and Bernhardt (1995) where a worker's current employer learns about the worker's ability perfectly after a single period of employment. Human capital accumulates deterministically with tenure.

In each period, worker i attains high (H) productive efficiency with probability $\theta_i = \theta_A \in \{\theta_g, \theta_r\}$ and low productive efficiency with probability $1 - \theta_i$. That is, a worker's ability type affects the probability of attaining high productive efficiency and a good worker attains high productive efficiency with a higher probability, i.e., $\theta_g > \theta_r$. The high or low production efficiency translates into different output realizations on different job levels. To be specific, worker i's output in period t if she is assigned to job j is

$$y_{it}^{j} = \begin{cases} (1+s_t)[z_H^j + f(x_{it})] & \text{with prob. } \theta_i \\ (1+s_t)[z_L^j + f(x_{it})] & \text{with prob. } (1-\theta_i), j \in \{l,m\}. \end{cases} \tag{2.1}$$

I assume $z_H^m > z_H^l > z_L^l > z_L^m > 0$. This set up captures that the manager job has greater returns to the high productive efficiency but the laborer job has greater returns to the low productive efficiency. It also follows the standard assumption in the job assignment literature as in Sattinger (1975) and Rosen (1982) that the manager job has greater marginal returns to an increase in productive efficiency from low to high.

In addition, I assume that a good worker is on average more productive on the manager job than on the laborer job but an ordinary worker is on average more productive on the laborer job than on the manager job. Let $E_A^j = \theta_A z_H^j + (1-\theta_A) z_L^j$, $A \in \{g, r\}$. Then $E_g^m > E_g^l$ and $E_r^l > E_r^m$. This set up means that a good worker has a comparative advantage producing as a manager but an ordinary worker has a comparative advantage producing as a laborer (although a good worker is always more productive than an ordinary worker on both jobs, i.e., $E_g^j > E_r^j, j \in \{l,m\}$). Therefore, firms have an incentive to (correctly) assign a good worker to a manager position and an ordinary worker to a laborer position.

Let q_{it}^{INC} denote the belief of an incumbent firm in period t that worker i is good based on her output history.[5] I refer to a worker's previous period's employer as the incumbent firm and all other firms as outside firms. Since the speed of learning on the lower level and the upper level job is the same, the belief that a worker is good is a function of

[5]Outside firms' beliefs about workers' ability types are based on workers' job assignments at their incumbent firms. I will discuss in more detail about how outside firms' beliefs are determined in equilibrium in the next section.

whether or not a worker attains high (low) productive efficiency only. That is, at which job level she has worked is irrelevant. In addition, given the binary ability types, only the total number of high (or low) productive efficiencies that a worker attains matters for the belief in a given period. Let h_i^{t-1} denote the total number of high productive efficiencies that worker i has attained up to period $t-1$. The expected output of worker i who is believed to be good with probability $q_{it}(h_i^{t-1})$ and who is assigned to job j in period t is

$$\begin{aligned} E[y_{it}^j | q_{it}^{INC}(h_i^{t-1})] &= (1+s_t)\{q_{it}^{INC}(h_i^{t-1})[\theta_g z_H^j + (1-\theta_g) z_L^j] \\ &\quad + [1-q_{it}^{INC}(h_i^{t-1})][\theta_r z_H^j + (1-\theta_r) z_L^j] + f(x_{it})\} \\ &= (1+s_t)\{q_{it}^{INC}(h_i^{t-1}) E_g^j + [1-q_{it}(h_i^{t-1})] E_r^j + f(x_{it})\}, j \in \{l, m\}. \end{aligned} \tag{2.2}$$

x_{it} is worker i's labor market experience up to period t. $f(\cdot)$ is the human capital accumulation function. Following Acemoglu and Pischke (1998), I assume f to be twice continuous differentiable, strictly increasing, concave with $f(0) \geq 0, \lim_{x \to 1} f'(x) = \infty$ and $\lim_{x \to \bar{t}} f'(x) = 0$ for some $2 \leq \bar{t} \leq T-1$. That is, human capital accumulates very fast when tenure is low. When tenure is higher than $\bar{t}$, human capital almost stops growing with additional tenure. The human capital accumulation function enters into the production function additively to the part of the output that is determined by workers' innate abilities.

$s_t = S > 0$ if a worker is employed by her previous period's employer in period t. $s_t = 0$ otherwise. s_t thus captures firm-specific human capital. Following Bernhardt (1995), I assume that once a worker leaves her previous employer, her previous employer becomes a new firm to her and cannot collect the firm-specific human capital anymore unless she comes back and works for her previous employer for another period. This assumption guarantees that in each period only one firm can collect the firm-specific human capital from a worker's productivity. This means that workers do not have an incentive to constantly change employers to enable the firm-specific human capital in multiple firms. In practice, it is possible that after an employee leaves a firm, the firm' policy, structure, or business practices change such that the previous firm-specific human capital is not applicable when this employee re-enters the firm.

There is a cutoff belief that a worker is good, q^*, such that the expected output on level l and that on level m are equal. Note that the equal-productivity cutoff at an outside firm is equal to the equal-productivity cutoff at an incumbent firm.[6] q^* solves $q^*E_g^l + (1-q^*)E_r^l = q^*E_g^m + (1-q^*)E_r^m$, $q^* = (E_r^l - E_r^m)/[(E_g^m - E_r^m) - (E_g^l - E_r^l)]$. Thus, if the belief that a worker is good in period t is above this cutoff level, she is expected to be more productive on the management job; otherwise, she is expected to be more productive on the laborer job.

Following Greenwald (1986), I assume that a small fraction, λ, of workers leave the incumbent firm for exogenous reasons in each period. The existence of exogenous job switchers alleviates the winner's curse problem as discussed in the previous section.[7] I consider equilibrium behavior when $\lambda \to 0$.

To keep the model tractable, I focus on parameterizations that satisfy the following two conditions.

(i) $q^* < q_{iT}(1)$. This condition says that if worker i attains only one high productive efficiency in any of the previous $T-1$ periods, she is more productive on the upper-level job in period T. Therefore, there are no demotions in equilibrium.

[6] In general, if the firm-specific human capital term is not multiplicative in the production function, the equal-productivity cutoff ability levels are different in an incumbent firm and in an outside firm.

[7] Note that the exogenous job-switching status is different in every period, i.e., an exogenous mover in this period might not be an exogenous mover in the next period. As I will show, in equilibrium, there is no turnover other than the exogenous job-movers. If the job-switching status is time-invariant, once a worker moves, she reveals her true job-switching type such that the firm can separate the exogenous movers from other workers perfectly. As a result, I would have the strong winner's curse for the non-exogenous movers at the beginning of period 3.

(ii) S is "large". In particular, I assume that S is large enough (the precise parameter restriction can be found in the Appendix) that, in period T, the incumbent firm has an incentive to assign a worker with only one high output realization up to period $T-1$ to the upper level job. This condition guarantees that a worker is promoted when a high productive efficiency is attained. It also means an incumbent firm's belief about a non-promoted worker's ability type based on realized outputs is the same as an outside firm's belief about this worker's ability type based on observed job assignments. It guarantees that there is promotion in every period and there is no distortion in firms' promotion decisions.

Firms and workers are risk neutral and discount the future with a common discounting factor $\beta < 1$. There is no cost to workers from changing firms or to firms from hiring or firing workers. Under these assumptions, long-term contracts are not necessary, so I consider equilibrium wages that are determined by spot-market contracts. At the beginning of each period, firms engage in a wage-setting game where they place wage "bids" in order to attract workers. That is, wages are promised before production begins in each period.

The timing of the events is the following. At the beginning of period 1, nature moves first to assign an ability type to each worker and this ability type is time invariant. Firms make period-1 job assignments and wage offers conditional on their prior beliefs' about a worker being good. Workers choose the firm with the highest wage offer to work at. At the end of period 1, incumbent firms privately observe workers' outputs. At the beginning of the next period, incumbent firms update their beliefs about workers' ability types and announce job assignment decisions for their previous period's employees. After outside firms observe these job assignment decisions, all firms make wage offers simultaneously. Workers privately learn about their job-switching types in this period and the exogenous movers depart. Workers then choose the firm with the highest wage offer to work at. If there are multiple firms offering the same highest wage, a worker chooses randomly among those highest-wage-offer firms but stays with her previous period's employer if her previous period's employer is one of the highest-wage-offer firms. Production then begins. At the end of period 2, workers' outputs are privately observed by their incumbent firms. This process repeats until date T.

Firms' strategies are sequences of job-assignment and wage-offer pairs. Let w_{it}^{j} denote the wage offer to worker i in period t. J_{it} denotes the job assignment to worker i in period t. In equilibrium, the incumbent firm anticipates that outside firms' behavior would be affected by its promotion decisions. The best response is to match the wage offer from an outside firm and extract the rent created by the firm-specific human capital. Therefore, the equilibrium wage offer from the incumbent firm is equal to the wage offer from the outside firm and there is no turnover in equilibrium except for the exogenous movers. A firm's strategy set is

$$\left\{J_{it}, w_{it}^{j}\right\}_{t}, j = J_{it} \in \{l,m\}, t \in \{1,\ldots,T\}.$$

Figure 2.1 shows the sequences of beliefs, job assignments, and wage offers after the incumbents observe the productive efficiencies of their own workers in each period.

2.4 The Analysis

In this section, I begin the analysis with a brief discussion about the equilibrium job assignment rules and wage offers when a worker's type is perfectly known. I then consider what happens given asymmetric information and discuss the relationship between non-promotion, negative signals, and wage dynamics.

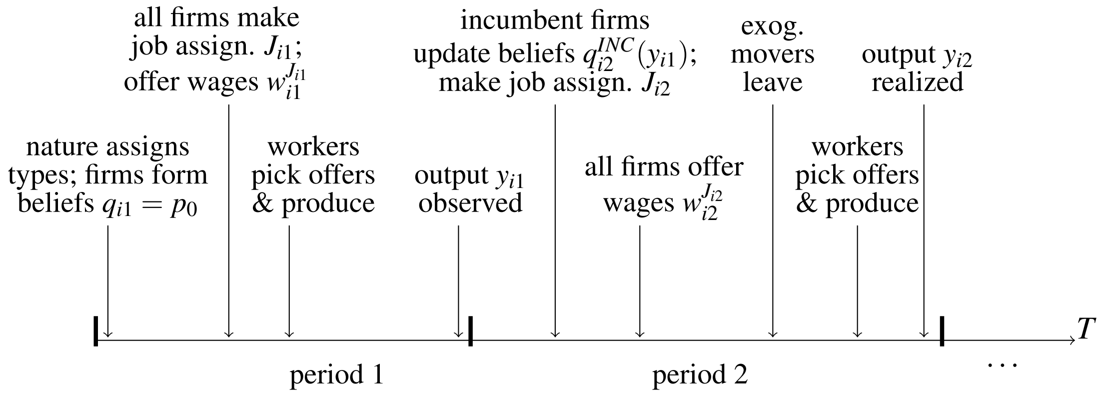

Figure 2.1: Timing of the job-assignment-wage-offer game.

2.4.1 A Full-information model with T periods

Under full information, workers' types are fully observed by all firms. Each firm's problem is to assign workers to the jobs that maximize the total discounted profit. Firms solve different problems when assigning workers who have worked for them in the previous period and those who have not. This is because firms can collect firm-specific human capital from the old workers in the current period but they have to train new workers.

Consider a firm's problem when it assigns workers who were employed by the firm in the previous period, i.e., when the firm makes job assignment decisions and wage offers as an incumbent firm. Let $\Pi_t^{INC}(\theta_i)$ denote the incumbent firm's profit in period t from employing a worker who has a probability θ_i of attaining the high productive efficiency. $\Pi_t^l(\theta_i)$ is the non-promotion profit and $\Pi_t^m(\theta_i)$ is the promotion profit from employing this worker. Then $\Pi_t^{INC}(\theta_i) = \max\left\{\Pi_t^l(\theta_i), \Pi_t^m(\theta_i)\right\}$,[8] where

[8]Since there is no turnover in equilibrium for non-exogenous movers and the probability of exogenous moving approaches zero, firm-tenure is equal to their total labor market experience in t. Also, I only have two job levels in the current set up. For those who are not promoted and who do not move to other firms, their labor market experience, firm-tenure, and job-tenure are the same. For the promoted workers in the period of promotion, their labor market experience, firm-tenure and job-tenure are also the same. However, the negative signals are embedded in the job-tenure. I will discuss how to extend the framework to distinguish between the three after I analyze the model.

$$\Pi_t^l(\theta_i) = (1+S)[\theta_i z_H^l + (1-\theta_i)z_L^l + f(t-1)] - w_{it}^l + \beta\Pi_{t+1}^{INC}(\theta_i), \tag{2.3}$$

$$\Pi_t^m(\theta_i) = (1+S)\,[\theta_i z_H^m + (1-\theta_i)z_L^m + f(t-1)] - w_{it}^m + \beta\Pi_{t+1}^{INC}(\theta_i). \tag{2.4}$$

w_t^j is the equilibrium wage that an incumbent firm offers, which is equal to outside firms' wage offer.[9] In period t, the incumbent firm chooses a job-assignment-wage-offer pair to maximize the total discounted profit given workers' types. If non-promotion is more profitable, the firm assigns the worker to the lower level job. Otherwise, the firm assigns the worker to the upper-level job. Due to the firm-specific human capital, an incumbent firm can potentially make positive profit by retaining an old worker.

When a firm makes decisions as an outside firm, i.e., when it considers wage bids for workers in other firms, its problem is characterized by a zero profit condition (due to free entry). Thus, the equilibrium wage in period t is equal to a worker's current period's productivity plus potential discounted future profits at an outside firm, $\Pi_{t+1}^{OUT}(\theta_i)$, i.e.,

$$w_{it}^j(\theta_i) = [\theta_i z_H^j + (1-\theta_i)z_L^j + f(t-1)] + \beta\Pi_{t+1}^{OUT}(\theta_i), j \in \{l,m\}, \theta_i \in \{\theta_g, \theta_r\}. \tag{2.5}$$

In the last period, since there are no future periods, a worker's wage is equal to her productivity in period T. In periods before T, a period-t "outside" firm would become an "incumbent" firm from period $t+1$ onward. This is because from period $t+1$ onward, a period-t outside firm starts to collect the firm-specific human capital as a period-t incumbent firm does. Since a period-t outside firm has the same information about a worker as a period-t incumbent firm does, the outside firm's expected future profit is equal to an incumbent firm's expected future profit from period $t+1$, i.e.,

[9]The expressions in 2.3 and 2.4 describe equilibrium behavior. The original firms' problem for this wage-bidding Bertrand competition is that the incubent firm and the outside firms choose their own wages holding other firms' wage bids as fixed. In equilibrium, those wages are equal. I omit the original expression to simplify the notation.

$\Pi_{t+1}^{INC}(\theta_i) = \Pi_{t+1}^{OUT}(\theta_i) = \Pi_{t+1}(\theta_i)$. Thus, we can substitute (2.5) into (2.3) and (2.4), and an incumbent's job-assignment problem simplifies to

$$\Pi_t^{INC}(\theta_i) = S \cdot \{ \max_{l,m} [\theta_i z_H^l + (1-\theta_i) z_L^l, \theta_i z_H^m + (1-\theta_i) z_L^m] + f(t-1)\}, \theta_i \in \{\theta_g, \theta_r\}.$$

Since a good worker has higher expected productivity on the upper-level job and an ordinary worker has higher expected productivity on the lower level job, the good worker should be assigned to the upper-level job and the ordinary worker should be assigned to the lower level job in each period. Also, since workers of the same type are ex ante identical, the equilibrium wages are only functions of workers' types, i.e., all good workers are paid the same wage in a certain period while all ordinary workers are paid another. I thus omit the individual subscript in the wage equations in the rest of this section.

I summarize the job assignment rules and equilibrium wages under full information in the following proposition. All proofs are provided in the appendix.

Proposition 3. *Suppose each worker's type is fully observable. Then the job assignment rules and the equilibrium wages satisfy (i) and (ii):*

(i) A good worker is assigned to the management level in every period and is paid $w_t^m(\theta_g) = [E_g^m + f(t-1)] + \beta S[E_g^m + f(t)]$ *in period* $1 \leq t < T$*; she is paid* $w_T^m(g) = E_g^m + f(T-1)$ *in period* T*;*

(ii) An ordinary worker is assigned to the laborer level in every period and is paid $w_t^l(\theta_r) = [E_r^l + f(t-1)] + \beta S[E_r^l + f(t)]$ *in period* $1 \leq t < T$*; she is paid* $w_{iT}^l(r) = E_r^l + f(T-1)$ *in period* T*.*

From Proposition 3, only the next period's productivity matters for the equilibrium wage. This is because the future profit is the extra economic rent that an incumbent firm can extract from collecting the benefit of one more period of firm-specific human capital compared to an outside firm. Thus, the rent is the next period's productivity multiplied by the firm-specific human capital factor S. This rent gives an outside firm an incentive to become an incumbent firm in the next period by bidding away a worker in the current period. Since firms are competing with each other over workers, all the economic rent is reflected in the equilibrium wage. The higher the firm-specific human capital is, the more a firm is willing to pay in anticipation of a higher rent from collecting the firm-specific part of the productivity.

Now let us consider how the timing of promotion is related to workers' wages. An ordinary worker's wage growth on the laborer job between two periods is $w^l_{t+1}(r) - w^l_t(r) = [f(t) - f(t-1)] + \beta S[f(t+1) - f(t)] > 0$. That is, a non-promoted worker's wage always increases with job tenure. A similar pattern is observed for good workers' wages. The reason is that under full information there is no learning with additional tenure, so workers' wages are determined solely by human capital accumulation which is non-decreasing with additional tenure. However, this prediction is inconsistent with the finding in Baker et al.(1994a) that non-promoted workers' wages first increase then decrease with tenure on the same job level for workers who spend a long time on the same job level. In the next section, I consider what happens when learning is asymmetric. I show that with asymmetric learning, non-promotion interacts with human capital accumulation and affects workers' wage dynamics through learning.

2.4.2 A Model with Asymmetric Information

With asymmetric information, incumbent firms observe outputs and update their beliefs about a worker being good and then make job assignment decisions and wage offers based on observed outputs. Outside firms observe workers' job assignments at the incumbent firms and update beliefs about workers' ability types. I focus on perfect Bayesian Equilibriums (PBE) of the model. That is, equilibrium beliefs are derived based on Bayes' rule given equilibrium strategies and equilibrium strategies are optimal for the incumbent firms, the outside firms, and the workers given the equilibrium beliefs.

Under the current framework, the incumbent firm never learns workers' true types. Furthermore, because of the binary output structure, when the incumbents make promotion announcements, they convey their private information about a worker's output in the current period completely under the parameterizations specified in the previous section. That is, whenever an incumbent firm observes the high-level productive efficiency, a worker is promoted; whenever an incumbent firm observes the low-level productive efficiency, a worker stays on the same job level (a promoted worker stays on the upper level job). In anticipating these equilibrium strategies, an outside firm believes that the high-level productive efficiency was attained at the laborer job when a promotion is observed; while a low-level productive efficiency was attained if a non-promotion is observed. It remains to check that under those beliefs the incumbents' strategies are indeed optimal.

The firms' problem is similar to the one under full information, i.e., incumbent firms choose a job-assignment-wage-offer pair to maximize total expected profit. The difference is that, since workers' types are unknown, workers' expected productivity is determined by the belief that a worker is good. Let $\Pi_t^{INC}(q_{it}^{INC})$ denote the incumbent firm's profit in period t from employing a worker who is believed (by the incumbent firm) to be good with probability q_{it}^{INC}. $\Pi_t^l(q_{it}^{INC})$ is the non-promotion profit and $\Pi_t^m(q_{it}^{INC})$ is the promotion profit from employing this worker. Then $\Pi_t^{INC}(q_{it}^{INC}) =$

$\max\left\{\Pi_t^l(q_{it}^{INC}), \Pi_t^m(q_{it}^{INC})\right\}$, where

$$\Pi_t^l(q_{it}^{INC}) = (1+S)[q_{it}^{INC}E_g^l + (1-q_{it}^{INC})E_r^l + f(t-1)] - w_{it}^l + \beta\Pi_{t+1}^{INC}(q_{it}^{INC}, l), \quad (2.6)$$

$$\Pi_t^m(q_{it}^{INC}) = (1+S)[q_{it}^{INC}E_g^l + (1-q_{it}^{INC})E_r^l + f(t-1)] - w_{it}^m + \beta\Pi_{t+1}^{INC}(q_{it}^{INC}, m). \quad (2.7)$$

$\Pi_{t+1}^{INC}(q_{it}^{INC}, j), j \in \{l, m\}$, is the incumbent's future expected profit given the belief in t that a worker is good with probability q_{it}^{INC} and the fact that the worker is assigned to job j in period t.

The outside firms' problem is characterized by a zero profit condition, so an outside firm is willing to bid above a worker's current period's expected productivity because it can collect future rents when it becomes an incumbent firm. Similar to the property in the full information equilibrium, in determining wages, only the next period's job assignment and productivity are relevant to the expected rents because an incumbent firm only collects one more period of firm-specific human capital compared to an outside firm.

Outside firms' wage bids are

$$w_{it}^l(q_{it}^{OUT}) = [q_{it}^{OUT}E_g^l + (1-q_{it}^{OUT})E_r^l + f(t-1)] + \beta\Pi_{t+1}^{OUT}(q_{it}^{OUT}, l), \quad (2.8)$$

$$w_{it}^m(q_{it}^{OUT}) = [q_{it}^{OUT}E_g^l + (1-q_{it}^{OUT})E_r^l + f(t-1)] + \beta\Pi_{t+1}^{OUT}(q_{it}^{OUT}, m), \quad (2.9)$$

where q_{it}^{OUT} denotes the outside firm's belief in period t that worker i is good; $\Pi_{t+1}^{OUT}(q_{it}^{OUT}, j)$ denotes the outside firm's future expected profit given the belief in t that a worker is good and the fact that the worker is assigned to job j in period t.

Note that under asymmetric information, an outside firm's belief that a worker is good is based on the incumbents' job assignment signals while the incumbent's belief is based on workers' output realizations. That is, an incumbent firm and an outside firm have different information sets about each worker. Thus, the outside firm's belief about the worker and the expected future rents from this worker might be different from the incumbent firm's. However, parameter restriction (ii) guarantees that the incumbent firm promotes a worker when a high output is observed. Therefore, at the time of promotion, the incumbent and the outside firms have the same information about a non-promoted worker, i.e., $q_{it}^{INC} = q_{it}^{OUT} = q_{it}, \Pi_{t+1}^{INC}(q_{it}^{INC}, j) = \Pi_{t+1}^{OUT}(q_{it}^{OUT}, j) = \Pi_{t+1}(q_{it}, j)$. Thus, an incumbent's job-assignment problem simplifies to

$$\begin{aligned}
\Pi_t[q_{it}(h_i^{t-1})] &= S \cdot \max_{l,m} \left\{ E[y_{it}^l | q_{it}(h_i^{t-1})], E[y_{it}^m | q_{it}(h_i^{t-1})] \right\} \\
&= S \cdot \max_{l,m} \left\{ q_{it}(h_i^{t-1}) E_g^l + [1 - q_{it}(h_i^{t-1})] E_r^l, q_{it}(h_i^{t-1}) E_g^m + [1 - q_{it}(h_i^{t-1})] E_r^m \right\} \\
&+ S \cdot f(t-1).
\end{aligned}$$

Therefore, the equilibrium strategy is that if the belief of a worker being good is greater than the threshold, i.e., $q_{it} \geq q^*$ (recall that q^* equates the expected productivity on the two job levels), the worker is promoted. Otherwise, the worker remains in the previous job level. Given parameter restriction (i), once a high productive efficiency is observed (for the first time), the belief of this worker being good would be above the threshold and thus this worker will be promoted by an incumbent firm, i.e., since $q_{it}(1) > q_{iT}(1) > q^*$, once a worker attains the high productive efficiency, she is assigned to the manager position and remains there independent of subsequent output realizations.

At an outside firm, given the observed job assignment history, the expected future profit to an outside firm is the same whether the worker is assigned to the manager position or the laborer position in the current period. Thus, an outside firm only considers the current period's expected output when deciding where to assign a worker. If an outside firm observes a promotion (non-promotion), it believes that this worker has produced high (low) output at the incumbent firm and thus she is more productive on the upper (lower) level job in the current period. Therefore, a promoted (non-promoted) worker at an incumbent firm is also assigned to the upper-level (lower-level) job at an outside firm.

Let $E_t[Z_{it}^j|q_{it}(h_i^{t-1})]$ denote the part of the expected productivity in t that is determined by workers' ability types given the observed information up to $t-1$, i.e., $E_t[Z_{it}^j|q_{it}(h_i^{t-1})] = q_{it}(h_i^{t-1})E_g^j + [1-q_{it}(h_i^{t-1})]E_r^j$. Let $E_t[Z_{it+1}^j|q_{it}(h_i^{t-1}), j]$ denote the part of the expected productivity in $t+1$ from period t's perspective that is determined by workers' ability types given the observed information up to $t-1$ and the job assignment in t. The expression for $E_t[Z_{it+1}^{J_{it+1}}|q_{it}(h_i^{t-1}), j]$ is given in the appendix.

I summarize the job assignment rules and equilibrium wages under asymmetric information in the following proposition.

Proposition 4. *Suppose workers' types are not observed but the incumbent firms can observe workers' outputs and the outside firms can observe workers' job assignments. Given the prior belief that a worker is good with probability p_0, the job assignment rules and the equilibrium wages satisfy (i) to (iv):*

(i) All workers are assigned to the lower level job in period 1.

(ii) Let t_i be the first period in which worker i produces the high output. Then worker i is assigned to the laborer position in each period $t, t \leq t_i$. Her wage in t is $w_{it}^l[q_{it}(0)] = \{E_t[Z_{it}^l|q_{it}(0)] + f(t-1)\} + \beta S\{E_t[Z_{it+1}^{J_{it+1}}|q_{it}(0), l] + f(t)\}$.

(iii) Worker i is assigned to the manager position in each period $t, t_i+1 \leq t < T$. Her wage in t is $w_{it}^m[q_{it_i+1}(1)] = \{E_t[Z_{it}^m|q_{it_i+1}(1)] + f(t-1)\} + \beta\{(1+S)E_t[Z_{it+1}^{J_{it+1}}|q_{it_i+1}(1), m] - E_t[Z_{it}^m|q_{it_i+1}(1)] + Sf(t)\}$.

(iv) In period T, if $t_i \leq T-1$, the worker is assigned to the manager job and is paid $w^m_{iT}[q_{it_i+1}(1)] = E[Z^m_{iT}|q_{it_i+1}(1)] + f(T-1)$; if $t_i > T-1$, she is assigned to the laborer job and is paid $w^l_{iT}[q_iT(0)] = E[Z^l_{iT}|q_{iT}(0)] + f(T-1)$.

Note that for workers who have attained the high productive efficiency before period $T-1$, outside firms' beliefs about their types stop updating once they are promoted because outside firms cannot infer their outputs from job assignments anymore (see (*iv* in Proposition 4). However, outside firms expect a promoted worker to produce either high or low in the next period because there is no winner's curse and incumbents do not observe workers' types perfectly. After a worker worked for a firm for one period, the firm starts to collect new information about this worker. Thus, the expected rent derived from employing this worker is the difference between the worker's expected productivity at an incumbent firm in the next period, $(1+S)\{E_t[Z^{J_{it+1}}_{it+1}|q_{it_i+1}(1), m] + f(t)\}$, and the worker's expected productivity at an outside firm in the next period, $E_t[Z^m_{it}|q_{it_i+1}(1)] + f(t)$.

Now, let us consider how the non-promotion wages change with job-tenure. Note that all workers with the same output history are ex ante identical (i.e., $q_{it}(0) = q_{kt}(0), i \neq k$). That is, all the non-promoted workers in period t are paid the same wage. In the following discussion, I omit the subscript i for individuals. The wage paid to a worker who is on level l for t periods (i.e. she has attained the low-level productive efficiency in the previous $t-1$ periods) thus is $w^l_t[q_t(0)]$, and the wage paid to a worker who is on level l for $t+1$ periods (i.e. she has attained the low-level of productive efficiency in the previous t periods), is $w^l_{t+1}[q_{t+1}(0)]$.

Corollary 4. *Under asymmetric information, there exists a t^*_1, $2 < t^*_1 \leq \bar{t}$, such that if the following conditions are satisfied, the non-promoted workers' wages increase in periods before t^*_1+1 and decrease in periods after t^*_1+1, i.e., for $t < t^*_1+1$, $w^l_t[q_t(0)] <$*

$w^l_{t+1}[q_{t+1}(0)]$*; for* $t > t^*_1+1$, $w^l_t[q_t(0)] > w^l_{t+1}[q_{t+1}(0)]$ *if*

$$f(t^*_1+1) - f(t^*_1) > [q_2(1) - q_3(1)][E^l_g - E^l_r], \text{ and} \tag{2.10}$$

$$f(t^*_1+2) - f(t^*_1+1) < [q_{\bar{t}}(1) - q_{\bar{t}+1}(1)][E^l_g - E^l_r].^{10} \tag{2.11}$$

Corollary 4 says that when human capital accumulation exceeds the negative learning about workers' ability with additional tenure, wages increase; when human capital grows little between two periods, the negative learning about workers' abilities lead to a wage decrease.

To see how the conditions in (2.10) and (2.11) guarantee the wage patterns described in Corollary 4, note that as additional low outputs are observed, the expectation that a worker is good decreases, i.e., $q_t(0) > q_{t+1}(0)$, for $t \geq 2$. Since $E^l_g > E^l_r$, the part of the wage that is related to workers' ability types, $E_t[Z^l_t|q_t(0)] = q_t(0)E^l_g + [1-q_t(0)]E^l_r$, decreases when firms put a smaller weight on the belief that a worker is good. Similarly, the forward expectation, $E_t[Z^{J_{t+1}}_{t+1}|q_t(0), l]$, also decreases in t, because it is less likely that a worker would produce high in the next period if she has produced more low outputs in the past. Since both expectations are bounded, if there is substantial human capital accumulation from period t to period $t+1$, the non-promotion wage increases. If the human capital accumulation from period t to period $t+1$ is sufficiently small, the non-promotion wage decreases. By construction, since human capital accumulates very fast when t approaches 1 and it almost stops growing after $\bar{t}$, there exists at least one period between period 2 (note that tenure in period 2 is equal to 1) and period $\bar{t}+1$ such that the non-promotion wage turns from increasing to decreasing.

Furthermore, the expectation about workers' types decreases at a decreasing speed and eventually approaches zero when t approaches ∞. Thus, the largest decrease in expectation is between period 2 and 3. On the other hand, human capital increases at a decreasing speed and eventually approaches zero when t approaches $\bar{t} < \infty$. Thus, the largest increase in human capital after period $t_1^* + 1$ is $f(t_1^* + 2) - f(t_1^* + 1)$. The condition in (2.10) guarantees that in periods before $t_1^* + 1$ the smallest human capital accumulation outweighs the largest expectation decrease. Thus, workers' wages increase before period $t_1^* + 1$. After that, the condition in (2.11) guarantees that the largest human capital accumulation is smaller than the smallest expectation decrease. Therefore, workers' wages decrease after period $t_1^* + 1$.[11]

As graphed in Baker et al. (1994a), for workers who are promoted from level 1 to level 2 within six years of tenure on level 1, their (real) wages prior to the promotion increase with each additional year of job-level tenure. For workers who are promoted after the sixth year, their wages prior to the promotion first increase then decrease with additional level-1 tenure. This empirical finding departs from Bernhardt's (1995) prediction that the non-promotion wages either increase or decrease monotonically with firm-level tenure but is captured in the above Corollary.

Using a similar argument, one can examine the wage-tenure relation for wages upon promotion.

Corollary 5. *Under asymmetric information, there exists a* t_2^*, $3 < t_2^* \leq \bar{t}$, *such that if the following conditions are satisfied, the promotion wages increase in periods before* $t_2^* + 1$ *and decrease in periods after* $t_2^* + 1$, *i.e., for* $3 \leq t < t_2^* + 1$, $w_t^m[q_t(1)] < w_{t+1}^m[q_{t+1}(1)]$; *for* $t > t_2^* + 1$, $w_t^m[q_t(1)] > w_{t+1}^m[q_{t+1}(1)]$ *if*

$$f(t_2^* + 1) - f(t_2^*) > \frac{1+S}{S}[q_4(2) - q_5(2)][E_g^l - E_r^l], \text{ and} \tag{2.12}$$

$$f(t_2^* + 2) - f(t_2^* + 1) < (1 - \beta)[q_{\bar{t}}(2) - q_{\bar{t}+1}(2)][E_g^l - E_r^l]. \tag{2.13}$$

[11] If the conditions in (2.10) and (2.11) are not satisfied, I still get the result that non-promoted workers' wages increase when tenure is low and their wages eventually fall (after $\bar{t} + 1$ for example). But I cannot guarantee that the non-promotion wages only turn once from increasing to decreasing.

Similar to the non-promotion wage, the wage in the period of promotion first increases then decreases with more time spent on the lower level job before promotion. Baker et al. (1994a) find that if a worker earns a promotion within four years on level 1, the wage that she earns upon promotion is higher than the wage paid to someone who is promoted in the previous period. On the other hand, if a worker earns a promotion after spending more than four years on level 1, the wage that she earns upon promotion is lower than the wage paid to someone who is promoted in the previous period.

Note that the condition in (2.12) is stronger than that in (2.10) since $q_t(2) - q_{t+1}(2) > q_t(1) - q_{t+1}(2) > q_t(0) - q_{t+1}(0)$. Therefore, it is possible that $t_2^* < t_1^*$, which means the promotion wage falls before the non-promotion wage does.

It is worth noticing that the signaling effect is embedded in job-tenure rather than firm-tenure. In the current set up, the level-l job-tenure is equal to firm-tenure before promotion. Suppose I were to extend the model to include a level below the laborer's level, call it the routine level, where workers' productivities do not vary with abilities and workers in this level are randomly selected into the laborer's level. If we compare a worker who has eleven years of firm tenure with four years on the routine job and seven years on the laborer's job to a worker who has ten years of firm-tenure with two years on the routine job and eight years on the laborer's job, if both workers are not promoted in the current period, the former would earn a higher wage than the latter although she has longer firm-tenure. The latter has a lower wage because she spends more time on the laborer's position. That is, the negative signal is associated with job-tenure rather than firm-tenure.

2.5 Data and Tests

In Baker et al. (1994a), the wage-job-tenure profile is shown using a raw plot. In this section, I use the same dataset that they have studied to estimate a tenure-wage equation controlling for other observables. I focus on the relationship between job-tenure and non-promotion wages as well as the relationship between job-tenure (before promotion) and promotion wages.

Table 2.1: Levels, Titles, and Education

Level	TITLE	HS	BS	MA	PHD	Total
1	AH	80	78	16	7	181
N=4,699		44.2	43.09	8.84	3.87	100
Pct.=29.63%	AJ	1,460	1,301	481	31	3,273
		44.61	39.75	14.7	0.95	100
	AK	766	402	77	0	1,245
		61.53	32.29	6.18	0	100
2	H	59	73	20	17	169
N=5,399		34.91	43.2	11.83	10.06	100
Pct.=34.05%	I	43	124	21	0	188
		22.87	65.96	11.17	0	100
	J	97	134	70	17	318
		30.5	42.14	22.01	5.35	100
	K	428	180	59	0	667
		64.17	26.99	8.85	0	100
	L	1,406	585	209	3	2,203
		63.82	26.55	9.49	0.14	100
	M	551	848	409	46	1,854
		29.72	45.74	22.06	2.48	100
3	F	21	80	22	0	123
N=5,759		17.07	65.04	17.89	0	100
Pct.=36.32%	G	2,327	1,905	1,056	105	5,393
		43.15	35.32	19.58	1.95	100
	SH	47	148	24	24	243
		19.34	60.91	9.88	9.88	100
	Total	7,285	5,858	2,464	250	15,857
		45.94	36.94	15.54	1.58	100

The dataset was constructed by George Baker, Michael Gibbs, and Bengt Holmstrom from the personnel records of a medium-sized US firm in the financial services industry. It contains detailed information on workers' demographic characteristics, tenure, subjective performance evaluation, and promotion history. In their seminal papers, Baker et al. (1994a;b) provide a thorough analysis of wage and career dynamics in this firm during a 20-year period from 1969 to 1988 using the full sample of managerial employees for a total of 68,437 employee-year data points. In this analysis, I restrict the sample to US white males to focus on the wage dynamics without concerning the gender-wage-gap. I also exclude demotion, which takes up 2%-3% of the sample. Since I use a one-year lag in calculating job-tenure, I exclude the entry cohort '69. I also exclude any data points with missing performance measure and I exclude years of schooling 15,17,19, and 20 since those years of schooling are hard to categorize with a degree measure. There are eight job levels, where level 8 is the CEO position.[12] I only look at workers on levels 1 to 3 since the promotion and wage dynamics on upper-level jobs might be very different from those on lower level jobs. Moreover, I want to focus my analysis on the same sample that generates the wage plots in Baker et al. (1994a). In those plots, they focus on the wage and job-tenure relations on level 1 through level 3. This sample selection procedure gives me a sample of 15,857 employee-year data points across three job levels.

Table 2.1 presents the 17 major job titles as specified in Baker et al. (1994a;b), grouped by job levels and interacted with education groups following DeVaro and Waldman (2012). Observations are roughly equally distributed across three job levels, with 30% from level 1, 34% from level 2, and 36% from level 3. There are 12 job titles on level 1 to 3 but there are one or two job titles that are the dominant job title on a particular level.

Table 2.2 presents the descriptive statistics. Supervisor subjective performance rat-

[12]See Baker et al. (1994a;b) for detail descriptions about how the job levels are constructed.

Table 2.2: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Real Salary (in 1988 dollar)	15857	46726.37	10483.57	20846.91	109890.10
ΔReal Salary	11859	1644.12	3047.10	-13522.83	30163.46
Δ% Real Salary	11859	0.04	0.07	-0.24	0.91
Age	15857	39.03	9.43	23.00	69.00
High School	15857	0.46	0.50	0.00	1.00
Bachelor	15857	0.37	0.48	0.00	1.00
Master	15857	0.16	0.36	0.00	1.00
PhD	15857	0.02	0.12	0.00	1.00
Performance Rating (t-1)	11859	1.95	0.71	1.00	5.00
Year at Company (t-1)	13372	3.74	3.51	0.00	18.00
Year at Level (t-1)	15086	2.60	2.56	0.00	17.00
Promotion	15857	0.15	0.36	0.00	1.00

ings are measured annually on a five-point scale where 1 denotes the best performance and 5 the worst. There are roughly equal numbers of employee-years in the three job levels. The average tenure at the firm is 3.7 years and the average tenure in the job level is 2.6 years. Workers on average spend 2.3 years on level 1 before being promoted to level 2. They spend a little longer, 2.6 years, on level 2 before being promoted to level 3. Around 15% of the employees are promoted in each sample year.

To test the job-tenure-wage profile, I consider the following wage equation.

$$w_{it} = \beta_0 + \beta_1 L_{it-1} + \beta_2 W_{i0} + \beta_3 X_{it-1} + \varepsilon_{it} \tag{2.14}$$

i indexes an individual and t indexes years. X_{it-1} is a vector of controls including age and age squared, education, and performance rating in the previous year. L_{it-1} is year at level in $t-1$ before promotion. Thus, it denotes the job-level tenure at the previous level for a just promoted worker and job-level tenure at the current level up to $t-1$ for a non-promoted worker. W_{i0} is worker i's first salary at the firm, which is a control for workers' initial characteristics (Belzil et al., 2012). Workers' initial characteristics need to be controlled for because the model predictions concern learning about ex ante identical workers. Since the initial wage carries rich information about an individual, I use it as a proxy for workers' initial heterogeneity.

Table 2.3: Wage-tenure profile before promotion (quadratic): level=1

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) FE
		Dependent Variable: Real Salary in 88 Dollars			
Yr. at Level	80.128	1,120.291**	771.463**	1,256.270**	1,002.114**
(t-1)	(53.370)	(111.222)	(124.559)	(214.247)	(127.373)
Yr at Level^2		-135.559**	-99.508**	-112.916**	-79.256**
(t-1)		(12.083)	(12.466)	(17.230)	(10.237)
Entry Salary	0.328**	0.328**	0.313**	0.353**	
	(0.014)	(0.014)	(0.014)	(0.024)	
Rating (t-1)				-2,333.563**	
				(244.979)	
Age (t-1)			1,170.840**	889.369**	1,905.166**
			(87.675)	(137.627)	(254.780)
Age^2(t-1)			-14.225**	-11.218**	-24.356**
			(1.071)	(1.593)	(2.885)
High School			-307.858	-827.032*	
			(250.931)	(354.364)	
Master			3,649.813**	2,721.757**	
			(349.193)	(612.507)	
PhD			2,863.510**	1,972.203	
			(890.531)	(1,145.488)	
Constant	31,559.206**	30,848.928**	8,778.494**	18,337.007**	3,235.066
	(341.818)	(346.468)	(1,642.439)	(2,709.431)	(5,608.402)
Observations	4,516	4,516	4,516	2,338	4,518
R-squared	0.147	0.166	0.223	0.237	0.954

Robust standard errors in parentheses
** $p<0.01$, * $p<0.05$

The sample that I use to test the non-promotion-wage-job-tenure relation is restricted to those individuals who are on the same job level in a particular year and I consider individuals on different job levels separately. The sample that I use to test the promotion-wage-job-tenure relation is restricted to those individuals who are just promoted in a particular year and I consider the individuals who are promoted from level 1 to level 2 and those who are promoted from level 2 to level 3 separately.

Table 2.3 looks at the relationship between job-tenure and non-promotion wages (in real terms) for those individuals on level 1. The table begins with the most parsimonious specification with only job-tenure in the previous period and the entry salary as the explanatory variables in Column (1). There is a positive relationship between wage and job-tenure but it is not statistically significant. Each column from (2) to (4) adds additional controls. Column (2) adds a quadratic term for job-tenure. Column (3) includes controls for age and education. Column (4) adds controls for performance rating. The hump-shaped wage-job-tenure relation remains after I control for performance rating. However, in the theoretical model workers' wages are not conditioned on their performance, so column (3) is a better test of the theory. In Column (5), I consider a fixed-effect model on individual level excluding performance ratings. The relationship between job-tenure and the non-promoted workers' wage persists regardless of model specifications. In particular, the non-promoted workers' wage first increases then starts to fall. These results match Baker et al.'s (1994a) wage plots very well.

Note that, the OLS models compare the average wages across individuals and the fixed-effect model captures within person wage dynamics. From the theoretical model, all the non-promoted workers in a given period are paid the same wage because they are identical in all other dimensions. In practice, workers differ in age, education, performances and other dimensions. Therefore, I control for other observables in the OLS models. However, there might still be other unobserved individual characteristics that are driving the results. So the estimates of the OLS models capture two effects. First, those individuals who stay on level 1 for a shorter period of time on average earn a higher wage than those who stay longer because the former are more able to earn a promotion earlier. Second, each individual has a smaller wage increase (and eventually wage decrease) with longer tenure on the same level. The theoretical model suggests that we should also observe the hump-shaped wage-tenure pattern when we compare within individuals and these predictions are supported by the estimates from the fixed-effect model.[13]

Table 2.4 examines the job-tenure-non-promotion-wage relation using job-tenure dummies instead of imposing the quadratic form on job-tenure in the wage equation. We can see that the non-promoted workers' wages increase with job-tenure then start to fall after four or five years. For example, from Column (1), with three year tenure on job 1, a worker's wage is $1,969 higher than the entry wage; with four year tenure on job 1, a worker's wage is only $1,544 higher than the entry wage, which means the worker's wage starts to fall in the fourth year. If a worker spends more than six years on level 1 without a promotion, her wage even falls below the entry level. These patterns remain after I control for age and education in Column (2). In Column (3), I add controls for performance rating. The wage pattern remains. Column (4) presents the fixed-effect estimates with job-tenure dummies. We can see that workers wages increase fast in the first five years on the job and starts to fall in the sixth year.

Table 2.5 and Table 2.6 repeat the analyses in Table 2.3 and Table 2.4 for workers who stay on level 2. The hump-shape wage-tenure profile is still evident in both specifications with either a quadratic term on year-at-level or job-tenure dummies. Column (4) in each table controls for age, education and performance rating. Column (5) considers a fixed-effect model. From the fixed-effect model, the non-promotion wage on level 2 starts to fall in year 5.

Note that I do not control for firm tenure for the non-promoted workers for two reasons. First, for individuals who are on level 1, their firm tenure is equal to their

[13]There is a discrepancy between the theory and the empirical specification in (2.14) that in the theory workers only differ in their time-to-promotion. That is, all workers with the same job assignment history are paid the same wage. As discussed in Gibbons and Waldman (1999a), one way to enrich the model is to include different observed education levels such that workers with the same job assignment history are paid differently.

Table 2.4: Wage-tenure profile before promotion (year dummies): level=1

	(1) OLS	(2) OLS	(3) OLS	(4) FE
Yr at Level(t-1)	*Dependent Variable: Real Salary in 88 Dollars*			
1	1,327.494**	850.815**	-677.111	1,193.591**
	(289.458)	(286.645)	(353.374)	(141.784)
2	1,603.746**	784.368*	407.647	2,138.382**
	(357.945)	(377.822)	(344.477)	(223.858)
3	1,969.184**	1,010.640*	800.016	2,845.665**
	(470.574)	(500.490)	(457.721)	(328.424)
4	1,544.368**	530.355	526.257	3,057.691**
	(544.806)	(569.671)	(535.426)	(388.907)
5	716.288	-267.627	188.531	3,398.000**
	(643.595)	(660.488)	(626.804)	(494.805)
6	-325.991	-1,109.076	-127.689	3,224.353**
	(780.660)	(813.315)	(753.214)	(601.316)
7	-944.832	-1,641.365	-850.610	3,142.963**
	(989.624)	(977.187)	(922.020)	(755.284)
8	-4,072.561**	-4,371.377**	-3,399.209**	2,132.144*
	(1,096.801)	(1,089.629)	(1,072.734)	(877.605)
9	-4,696.114**	-5,248.299**	-4,232.494**	2,884.836**
	(1,485.813)	(1,480.712)	(1,419.875)	(1,084.924)
10	-6,822.486**	-6,992.827**	-5,512.765**	3,278.206*
	(1,876.262)	(1,573.646)	(1,565.759)	(1,424.512)
11	-4,745.591	-5,859.753*	-4,843.347	2,259.449
	(3,485.411)	(2,812.724)	(2,894.402)	(1,975.629)
12	-6,727.453**	-6,466.210**	-5,047.442**	1,521.800
	(807.280)	(1,284.532)	(1,300.991)	(1,391.815)
13	-7,321.807**	-6,495.907**	-5,126.091**	1,897.945
	(913.263)	(1,207.943)	(1,206.696)	(1,656.205)
Entry Salary	0.327**	0.313**	0.354**	*No*
	(0.014)	(0.014)	(0.024)	*No*
Age	*No*	*Yes*	*Yes*	*Yes*
Education	*No*	*Yes*	*Yes*	*No*
Rating	*No*	*No*	*Yes*	*No*
Constant	38,402.605**	13,989.254**	24,168.928**	7,151.725
	(175.917)	(1,729.504)	(3,103.598)	(5,009.659)
Observations	4,518	4,518	2,338	4,518
R-squared	0.026	0.097	0.107	0.954

Robust standard errors in parentheses
** p<0.01, * p<0.05

Table 2.5: Wage-tenure profile before promotion(quadratic): level=2

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) FE
	Dependent Variable: Real Salary in 88 Dollars				
Yr. at Level	-20.267	1,084.606**	733.951**	913.175**	454.667**
(t-1)	(34.432)	(128.834)	(146.204)	(166.053)	(88.055)
Yr at Level^2		-88.054**	-59.899**	-46.615**	-50.066**
(t-1)		(9.830)	(10.967)	(13.778)	(6.686)
Entry Salary	0.308**	0.338**	0.317**	0.373**	
	(0.013)	(0.013)	(0.013)	(0.016)	
Rating (t-1)				-2,361.349**	
				(181.537)	
Age (t-1)			1,179.609**	1,117.072**	2,388.337**
			(91.296)	(101.540)	(220.122)
Age^2(t-1)			-13.988**	-13.186**	-25.694**
			(1.078)	(1.179)	(2.488)
High School			-801.273**	-791.539**	
			(265.289)	(283.761)	
Master			3,055.688**	2,475.624**	
			(342.154)	(378.475)	
PhD			1,960.592*	1,313.417	
			(944.297)	(1,115.009)	
Constant	37,490.788**	35,143.455**	12,620.226**	16,757.512**	-7,025.540
	(317.198)	(425.766)	(1,764.041)	(1,980.936)	(4,636.120)
Observations	4,574	4,574	4,574	3,779	5,141
R-squared	0.111	0.127	0.174	0.199	0.930

Robust standard errors in parentheses
** $p<0.01$, * $p<0.05$

job tenure. Second, for a non-promoted worker on level 2, conditional on firm tenure, those who have longer level-2 tenure in general should have shorter level-1 tenure. If more able workers are promoted earlier on level 1, those individuals who have longer level-2 tenure should earn a higher wage. On the other hand, long tenure on level 2 sends a negative signal. So, holding firm tenure fixed, the theory predicts that wages can either increase or decrease with job tenure on level 2. The current specification without controlling for firm tenure examines the average effect of level-2 tenure on wages allowing individuals to have different years of tenure on level 1.

Table 2.6: Wage-tenure profile before promotion(year dummies): level=2

	(1) OLS	(2) OLS	(3) OLS	(4) FE
Yr at Level(t-1)	*Dependent Variable: Real Salary in 88 Dollars*			
1	3,642.644**	2,721.035**	-	2,237.414**
	(596.568)	(544.034)	-	(335.115)
2	4,056.761**	2,868.391**	778.697*	2,414.836**
	(620.571)	(586.657)	(330.145)	(364.138)
3	4,151.822**	2,798.967**	1,098.994**	2,829.843**
	(656.452)	(633.962)	(395.853)	(378.361)
4	4,976.847**	3,474.826**	2,227.494**	2,888.903**
	(706.598)	(692.060)	(484.368)	(409.393)
5	5,318.720**	3,808.978**	2,861.933**	2,760.834**
	(764.215)	(759.700)	(573.743)	(439.888)
6	5,472.409**	3,956.963**	2,869.221**	2,241.372**
	(873.991)	(861.200)	(681.666)	(496.213)
7	6,864.477**	5,570.411**	4,544.755**	2,363.988**
	(1,089.764)	(1,059.840)	(909.445)	(622.746)
8	4,548.405**	3,488.197**	2,876.764**	1,170.001
	(1,075.865)	(1,043.605)	(913.311)	(667.476)
9	4,452.337**	3,531.212**	2,737.820**	-37.641
	(1,073.098)	(1,079.420)	(913.661)	(729.459)
10	2,998.101**	2,133.263	1,864.390	-1,141.041
	(1,162.290)	(1,244.446)	(1,069.091)	(815.523)
11	4,278.598**	3,146.789*	2,939.162*	-797.063
	(1,457.885)	(1,494.909)	(1,360.042)	(917.085)
12	5,437.926**	4,927.277*	4,471.845*	-815.763
	(2,057.511)	(2,145.260)	(2,036.756)	(1,153.356)
13	1,343.494	876.404	4,195.668**	-488.140
	(711.369)	(657.982)	(1,034.647)	(454.285)
Entry Salary	0.345**	0.324**	0.373**	*No*
	(0.013)	(0.013)	(0.016)	*No*
Age	*No*	*Yes*	*Yes*	*Yes*
Education	*No*	*Yes*	*Yes*	*No*
Rating	*No*	*No*	*Yes*	*No*
Constant	42,978.276**	17,113.012**	22,109.807**	-7,541.936
	(570.767)	(1,696.482)	(1,956.427)	(4,714.525)
Observations	5,141	5,141	4,235	5,141
R-squared	0.011	0.069	0.079	0.932

Robust standard errors in parentheses
** p<0.01, * p<0.05

Table 2.7 examines the average wages in the year of promotion for workers with different job-tenure on the lower level job before promotion. Columns (1) and (2) consider the promotion wage when a worker is promoted from level 1 to level 2. From Column (1), the promotion wage first increases then decreases with tenure on the job before promotion but the estimates are not statistically significant. From Column (2), the promotion wage increases when tenure on the previous job is low and decreases with tenure on the previous job is high but the relationship flips signs several times. From Corollary 5, the promotion wage may change signs multiple times if the conditions in (2.13) and (2.13) are not satisfied. Thus, these wage patterns are not inconsistent with the model.

Columns (3) and (4) in Table 2.7 look at the promotion wage when a worker is promoted from level 2 to level 3. Overall, if a worker spends less time on level 2 before she is promoted, she earns a higher wage upon promotion. We do not observe the promotion wage increase when the job-tenure on level 2 is low. As I discussed in the theoretical analysis, under certain parameterization, it is possible that the promotion wage falls when the non-promotion wage is still increasing because the promotion wage is more sensitive to learning.

In summary, the empirical evidence supports the model's prediction about the non-promoted workers' wages and job-tenure. Job-tenure not only affects workers' human capital levels but also carries rich information about workers' unobserved ability. As discussed in Gibbs (1995), job tenure can be used as a proxy for workers' unobserved ability.

Table 2.7: Promotion Wage and Job Tenure Before Promotion

	level 1 to level 2		level 2 to level 3	
	(1)	(2)	(3)	(4)
	Dependent Variable: Real Salary in 88 Dollars			
Year at Level (t-1)	339.530		-380.067	
	(448.027)		(465.726)	
Year at Level Sq. (t-1)	-21.349		27.759	
	(55.417)		(46.083)	
Year at Level (t-1)				
2		-143.129		-856.522
		(509.929)		(647.654)
3		-162.140		39.202
		(656.884)		(869.102)
4		1,252.088		-1,354.274
		(898.077)		(1,098.808)
5		814.904		606.428
		(1,310.348)		(1,646.929)
6		2,086.035		-1,271.724
		(1,401.184)		(1,590.442)
7		92.246		-4,181.807*
		(2,003.750)		(1,834.913)
8		1,140.458		5,537.766
		(1,817.991)		(3,156.209)
9		-2,937.925**		-4,234.630**
		(1,030.795)		(1,420.881)
10		2,671.382		-3,086.256*
		(5,183.328)		(1,243.448)
11		-7,249.871**		-
		(661.319)		-
12		-		5,727.921**
		-		(1,374.278)
Entry Salary	0.335**	0.336**	0.315**	0.312**
	(0.026)	(0.026)	(0.030)	(0.031)
Age	*Yes*	*Yes*	*Yes*	*Yes*
Education	*Yes*	*Yes*	*Yes*	*Yes*
Constant	17,521.364**	17,164.737**	26,881.535**	26,979.739**
	(3,334.137)	(3,448.835)	(4,731.407)	(4,881.879)
Observations	1,170	1,170	910	910
R-squared	0.173	0.178	0.168	0.181

Robust standard errors in parentheses
** p<0.01, * p<0.05

2.6 Conclusion

This essay develops a theoretical framework to explore the negative signals associated with non-promotion. It contributes to the literature in three different ways. First, it fills a gap in the learning literature by capturing the negative signaling role of non-promotion. Second, this essay emphasizes the relationship between wages and job-level tenure instead of firm-level tenure. Third, it provides a systematic explanation for a set of empirical findings that are not well captured in existing models. My model shows that non-promoted workers' wages decrease when they spend a long time on the same job level, while their wages increase in the early years on the job. The empirical tests show clear evidence for a hump-shaped wage-job-tenure profile for the non-promoted workers. These results suggest that, besides determining workers' levels of human capital, job tenure carries additional information about individuals' unobserved ability. The trade-off between negative learning and positive human capital accumulation associated with additional tenure shapes the wage-job-tenure profile.

There are a number of ways to extend the model. First, there is no turnover in this model except for the exogenous movers. If workers are fully aware of the negative signals associated with non-promotion, they may choose to leave before the negative signals about their abilities are revealed through non-promotion. Second, I can allow for heterogeneity in the initial human capital stock by incorporating education. The strength of the negative signals is expected to be different for workers with different initial levels of human capital.

CHAPTER 3

FLATTENING FIRMS AND WAGE DISTRIBUTION

3.1 Introduction

Firms organize their workers into hierarchies to carry out production (Williamson, 1967; Calvo and Wellisz, 1979; Rosen, 1982). The hierarchical structure of firms is closely related to individuals' career and wage dynamics. In the past thirty years, firms became flatter, i.e., they got rid of some layers in their corporate hierarchy. This delayering trend is well documented in numerous studies using large-scale firm-level data sets from several developed countries (Colombo and Delmastro, 1999;2008; Rajan and Wulf, 2006; Caliendo et al., 2012). Widely identified and most convincing causes of this trend are the increased competition in the product market (Bloom et al., 2010; Guadalupe and Wulf, 2010), the improvement in corporate governance, and the advancement of information technology (Garicano, 2000; Bresnahan et al., 2002).

While the causes of this delayering trend are extensively studied, the consequences of this change, especially the impact on individuals' wages and the wage distribution within a firm, are not well explored.[1] In this essay, I build a model to explain two empirical regularities found in the recent literature. First, after firms delayer, wages at all levels increase (Bauer and Bender, 2001; Rajan and Wulf, 2006; Caliendo et al., 2012). Second, after delayering, the wage distribution becomes more unequal (Bauer and Bender, 2001).

To explain these wage patterns associated with delayering, I consider a job-

[1]Numerous studies in the management and human resource literature explore the relationship between firm delayering and subsequent firm performance. The conclusion is mixed. For example, Carzo and Yanouzas (1969) find that tall organizations are more profitable, while Shaw and Schneier (1993), Cristini et al. (2003), and Kuhn (2011) find that the opposite is true.

assignment model with slot constraints and asymmetric information. In this set up, workers compete for a single position at the upper level. Outside firms do not observe workers' outputs but can make inferences about workers' abilities using promotions as signals (Waldman, 1984a). In addition, I assume that workers sometimes leave their positions for exogenous reasons following Greenwald (1986). The main intuition in this model is that, holding firm size constant, the contestants' pool is larger in a flatter firm than in the firm with more layers. Therefore, winning a promotion in a flatter firm sends a more positive signal about the winner's ability.[2] In addition, losing a promotion in a larger contest sends a more positive signal about losers' abilities as well. This is because losing a larger contest does not necessarily mean that a worker is incompetent but rather that she is not the best among many workers. Thus, after delayering, wages at all levels go up. In addition, because the market expectation about the winner's ability goes up faster than the market expectation about losers' abilities, the wage gap between the winner and the losers widens after delayering.

Three key features of the model are essential to the resulting wage patterns. First, there is asymmetric learning among firms about workers' abilities. With asymmetric learning, workers' wages are determined by their "market value", which is based on their job assignments. Second, the upper level positions are characterized by a slot constraint. Without a slot constraint, after delayering, there would be an upward adjustment in the expectation about the workers' ability at the low level job and a non-upward adjustment in the expectation about the workers' ability at the upper-level job so that the wage distribution would become more equal.[3] Third, due to the existence of exogenous job

[2]This argument is related to discussions in Prendergast (1999) and Waldman (2013).

[3]For example, suppose there are three types of workers with high, regular, or low ability working for a three-layer firm. The high ability individual is at the top, the regular ability individual is in the middle, and the low ability individual is at the bottom. After delayering, since the regular worker is not good enough for the top level job, she would join the low ability worker on the low level job. As a result, the average ability at the top is unchanged while the average ability at the bottom rises. Suppose there is a continuum of workers' types, some middle level workers would join the top level job, and some middle level workers would join the bottom level job. As a result, the average ability at the top falls while the

movers, my model alleviates the “winner’s curse” problem that is common in models with asymmetric information (Milgrom and Oster, 1987). Without exogenous movers, if a worker is not promoted, the market would not want to offer a wage that is higher than the lowest expected productivity. If a worker’s previous employer learns her ability perfectly after one period as assumed in Waldman (1984a) and Zabojnik and Bernhardt (1995), this lowest expected productivity does not change and so delayering would not affect the wage for low level workers.

This essay contributes to the literature in several different ways. First, it contributes to the delayering literature by exploring the effects of delayering on wage changes. Second, it contributes to the job assignment literature by considering how firms’ organization structure affects wages. This essay also captures several empirical findings that are not well explained in the existing literature.

The organization of this essay is as follows. In Section 2, I review related literature. Section 3 discusses the model set up. In Section 4, I first analyze a model with two layers, and then compare the results to a three-layer model. Section 5 concludes.

3.2 Related Literature

As I discussed in the introduction, most of the empirical literature on delayering has focused on the causes of delayering (Garicano, 2000; Bresnahan et al., 2002; Bloom et al., 2010; Guadalupe and Wulf, 2010). Several recent studies have documented some new wage patterns related to firm delayering. For example, using data from more than 300 US firms from 1986-1998, Rajan and Wulf (2006) find that after delayering, the division

average ability at the bottom rises. Since workers’ wages are largely attached to their expected ability (productivity), wages in the firm become more equal.

managers are paid more in salary and bonus after controlling for firm size and firm fixed effects. Internationally, using a comprehensive sample of French manufacturing firms, Caliendo et al. (2012) find that after delayering, wages at all levels of the firm increases. In an earlier study using a nationally representative linked employer-employee panel dataset from Germany, Bauer and Bender (2001) find that average wages increase after firms delayer. The wage distribution also becomes more unequal.

Two existing theories have addressed the relationship between delayering and wage distributions within firms. The first stream of literature is based on the command-and-control argument. Qian (1994) extends Calvo and Wellisz (1979) by endogenizing the number of layers in a hierarchy. The main mechanism is that since the entrepreneur's attention is limited, the further down a worker is in the hierarchy, the looser the control is and the lower the worker's effort is. Since the optimal number of layers decreases as the capital stock shrinks, workers who remain in the same position relative to the bottom receive a higher wage due to an increase in control with a shorter chain-of-command and workers who remain in the same position relative to the top receive a lower wage due to a decrease in monitoring with a larger span. That is, delayering is associated with a wage increase at the lower level and a wage decrease at the top. This prediction is inconsistent with the recent findings in the delayering literature and the more established stylized fact that the CEO-to-average-wage ratio has increased dramatically over the past thirty years.[4]

The other strand of literature takes the knowledge hierarchy approach (Garicano, 2000). Caliendo and Rossi-Hansberg (2012) study a model in which firms eliminate layers in response to negative demand shocks. Since the total knowledge for production is unchanged, as the number of layers decreases, the knowledge and thus wages in all

[4]For example, Murphy and Zabojnik (2004) find that from 1970 to 2000, the ratio of CEO cash compensation to average pay for production workers increased from 25 to 2000.

pre-existing layers rise. In another study, Garicano and Rossi-Hansberg (2006) consider a model with homogeneous firms and heterogeneous workers. In their model, as the cost of acquiring knowledge decreases, fewer layers are preferred and managers acquire more knowledge. Since the knowledge increases more at the top, the overall wage inequality increases. While Caliendo and Rossi-Hansberg (2012) provide an explanation for why wages at all levels increase after delayering, they do not explain why wage inequality rises after delayering. On the other hand, Garicano and Rossi-Hansberg (2006) explain why wage inequality rises after delayering, but they do not explain why wages at all levels increase after delayering. My analysis captures both wage patterns associated with delayering under a single theoretical framework.

My model is built on the promotion-as-signal approach found initially in Waldman (1984a).[5] Using this approach, Waldman (1984a) and various extensions capture many stylized facts about wage and promotion dynamics, such as large wage increases upon promotion (Bernhardt, 1995) and the wage-and-firm-size effect (Zabojnik and Bernhardt, 2001), etc. This paper adds to the promotion-as-signal literature by looking at the effect of delayering on wages.

This model's set up is closely related to Zabojnik and Bernhardt (2001). Zabojnik and Bernhardt (2001) consider a model with slot constraints to explore the firm-wage-size relation. The main focus in their model is on how promotion can induce optimal human capital investment. In this paper, on the other hand, I explore the implications of slot constraints and firms' hierarchical structural changes on wage distributions. I also show that, different from the prediction in Zabojnik and Bernhardt (2001), more

[5]The promotion-as-signal approach assumes asymmetric learning in the labor market which means a worker's current employer knows more about the worker's true ability than outside firms do. A competing modeling framework is to assume symmetric learning where all firms have the same information about a worker's type (Harris and Holmstrom, 1982; Gibbons and Waldman, 1999). Although both frameworks can explain various stylized facts concerning wage and promotion dynamics inside firms, many recent empirical studies have found evidence in favor of the asymmetric learning framework (Pinkston, 2009; DeVaro and Waldman, 2012; Kahn, 2013). I thus adopt the asymmetric learning framework in this essay.

efficient firms are not necessarily larger when we take into account the internal structure of the firm.

3.3 The Model

In this section I set up a two-period model. There are $F(>2)$ identical firms in the market. Each firm hires n risk-neutral young workers in period 1 (and firms can choose to vary their sizes in period 2). Different jobs in a firm have different production efficiencies, denoted by V. The production efficiency translates workers' ability into output. Workers are ex ante identical and have two-period careers. I refer to an individual in her first work period as young, and those who are in their second work period as old. Worker i's ability, θ_i, is drawn from a uniform distribution on $[\theta_L, \theta_H]$.

An individual with ability θ who is assigned to a job with production efficiency V produces $s_t V\theta$ units of output in period t. $s_t = S > 1$ for an old worker who remains at her previous period's employer and $s_t = 1$ for a young worker or for an old worker who just starts to work for a new firm. s_t thus captures firm specific human capital. The total production at each firm is the sum of each worker's output. I refer to a worker's previous period's employer as the incumbent firm and all other firms as outside firms.

I assume there is over-supply of labor in the economy meaning that there are more than $n \cdot F$ workers. All the workers who are not hired by the firms stay self-employed. Following Waldman and Zax (2013), I assume that there is learning-by-doing in self-employment: a worker in her first period of self-employment produces $\bar{U}_1$ and a worker in her second period of self-employment produces $\bar{U}_2$, where $\bar{U}_2 > \bar{U}_1$. If a worker works for a firm in period 1 and becomes self-employed in period 2, she can only produce $\bar{U}_1$ in period 2. I assume $SV\theta_L > \bar{U}_1$, which means working for a firm in period 2 is better than self-employment if a worker has worked for a firm in period 1. This condition guarantees that workers do not change from working for a firm to self-employment between two periods. Furthermore, I assume $\bar{U}_1 + \bar{U}_2 > (1+S)VE(\theta)$, where $E(\theta)$ is the unconditional mean of a worker's ability. This condition guarantees that a firm hires a finite number of workers. This is because if firms hire an infinite number of workers, the total expected productivity and thus wage pay of each worker is $(1+S)VE(\theta)$. If this value is smaller than workers' total income from self-employment for two periods, $\bar{U}_1 + \bar{U}_2$, no worker would choose employment at a firm over self-employment. This regulates the firms not to hire an infinite number of workers.

Following Greenwald (1986), I assume a small probability, λ, that workers leave the firm for exogenous reasons. I consider the equilibrium behavior where $\lambda \to 0$. In addition, I assume that the firm-specific human capital is sufficiently large that an incumbent worker is always more productive than an outside worker.

The timing of the events is the following. At the beginning of period 1, nature assigns an ability type to each worker. Firms decide the optimal number of young workers to hire and offer wages accordingly. Workers choose the firm with the highest wage offer to work at. At the end of period 1, incumbent firms privately observe workers' outputs. At the beginning of period 2, incumbent firms update their beliefs about workers' types and one of the young workers is chosen to fill the upper level position if there is a vacancy at the upper level. Outside firms observe the incumbent firms' job assignment decisions and update their belief about workers' types. All firms then make wage offers simultaneously. Workers privately learn about their job-switching types and the exogenous movers depart. Workers then choose the firm with the highest wage offer to work at for period 2. If there are multiple firms offering the same highest wage, a worker chooses randomly among those highest-wage-offer firms but stays with her incumbent firm if her incumbent firm is one of the highest-wage-offer firms.

3.4 The Analysis

In equilibrium, firms assign workers to jobs and offer wages accordingly. I focus on the perfect Bayes-Nash equilibrium of this game. I first discuss equilibrium behavior in a two-layer firm and then I discuss and compare equilibrium behavior in a three-layer firm. Both models are solved by backward induction.

3.4.1 A two-layer model

Let us first consider the wage setting process in a two-layer firm. A two-layer firm consists of a CEO position (E) at the upper level and a number of laborer positions (L) at the lower level. Let $V^j, j \in \{E, L\}$ denote the production efficiency in job j. I assume $V^E > V^L$, which means the upper-level jobs have greater marginal returns to ability as in Sattinger (1975) and Rosen (1982).

In period 1, n young workers are hired into the laborer positions in each firm. In period 2, when the workers are old, one of them will be chosen to fill the CEO position (the CEO position remains unfilled in the first period). Since $V^E > V^L$, firms always have an incentive to assign the highest ability worker to the CEO position because this worker has the largest output increase if placed in the CEO position instead of a laborer position. In addition, due to the firm specific human capital, an incumbent firm can extract the highest rent from placing the most able worker in the CEO position.

In period 2, after observing a young worker's job assignment at her incumbent firm, outside firms form expectations about the worker's ability. Denote the expected ability of the promoted worker (i.e., the tournament winner) $\theta^E(n)$ and the expected ability of a non-promoted worker (i.e., a tournament loser) $\theta^L(n)$. Then $\theta^E(n) = E(\theta_i | i$ is the best among n laborers) and $\theta^L(n) = E(\theta_i | i$ is not the best among n laborers). Note that the expected ability is a function of the number of contestants in the promotion tournament only. $W^E(V,n)$ denotes the wage of the tournament winner who is promoted to the CEO position. $W^L(V,n)$ denotes the wage of the laborers who do not win the tournament. Proposition 5 describes the wages for the old workers in a two-layer firm. All proofs are reserved for the Appendix.

Proposition 5. *In equilibrium, an old worker's wage is equal to her expected productivity at an outside firm given her job assignment, i.e.,* $W^E(V,n) = V^L\theta^E(n)$ *and* $W^L(V,n) = V^L\theta^L(n)$. *The wage difference between the CEO and the laborers is* $\Delta W = W^E(V,n) - W^L(V,n)$, *which increases in the total number of workers in the laborer level n.*

Proposition 5 says that the old workers' wages are equal to their expected productivities at an outside firm, which are determined by their expected abilities as well as the production efficiency. Note that the wage paid to a tournament winner is evaluated at the laborer level. This is because firms are slot constrained at the upper level. In addition, due to the firm-specific human capital, an "insider" is more productive than an "outsider". Thus, firms do not replace an incumbent with an outside worker. Therefore, firms can only offer a wage that is consistent with assigning an outside worker to the lower level job.

Consider how the wage inequality, $\Delta W = V^L[\theta^E(n) - \theta^L(n)]$, changes with the total number of young workers hired into the laborer positions in period 1. When the pool of young workers (i.e. contestants) becomes larger, the winner's expected ability rises because the winner of a larger contest signals more strongly about her ability. On the other hand, the loser's expected ability also rises since not being the best among n laborers is a less bad outcome if n is larger.

To better illustrate this point, let us compare a two-person contest versus an n-person contest. Ex ante, the two groups of workers have the same expected ability. Ex post, since the expected ability of the winner from the two-person contest is above average, the loser's expected ability should be below average. That is, the "penalty" on losers' expected ability is fully absorbed by this one "loser". On the other hand, although the expected ability of the winner from the n-person contest is larger than that from the two-person contest, the penalty on losers' expected ability is averaged across many losers. The more the workers are in the promotion contest, the lower the penalty is on each losers' expected ability, and the closer the losers' expected ability is to the unconditional mean ability. In the extreme, when the size of the contestants' pool approaches infinity, the non-promoted workers' expected ability approaches the unconditional mean. Therefore, the expected ability of the losers from a larger contest is larger than the expected ability of the losers from a smaller contest.[6] When workers' ability is uniformly distributed, the expected ability for a tournament winner grows faster than the expected ability for a tournament loser as the size of the contestants' pool increases. To see this, note that $\theta^E(n) - \theta^L(n) = \frac{\theta_H - \theta_L}{2} \cdot \frac{1}{1+1/n}$. Thus, the wage gap between the winner and the losers widens monotonically with the size of the contestants' pool, i.e., $\partial W / \partial n > 0$.

Now, let us consider how firms choose the optimal number of young laborers to hire in period 1. Let n denote the number of young laborers. W_Y^L denote the first period's

[6]Waldman (2013) makes a similar argument and he illustrates this argument by considering a binomial case where workers are either good or bad. See footnote 19 in Waldman (2013).

wage for young laborers, which is a function of n. The firms' problem is the following.

$$\begin{aligned} \underset{n}{Max} \quad & nV^L E(\theta) - nW_Y^L(n) + [SV^E \theta^E(n) - W^E(V,n)] & (3.1) \\ & +(n-1)[SV^L \theta^L(n) - W^L(V,n)] \\ s.t. \quad & W_Y^L(n) + [\tfrac{1}{n} W^E(V,n) + \tfrac{n-1}{n} W^L(V,n)] \geq \bar{U}_1 + \bar{U}_2, & (3.2) \end{aligned}$$

(3.2) is workers' participation constraint that a worker is better off choosing working for a firm for two periods than staying self-employed for two periods. That is, the overall lifetime income should be at least the amount of what the self-employment pays. In equilibrium, the participation constraint binds. The first-order condition to the above problem is

$$S(V^E - V^L)\frac{\theta_H - \theta_L}{(n+1)^2} = (\bar{U}_1 + \bar{U}_2) - (1+S)V^L E(\theta). \quad (3.3)$$

From (3.3), if we hold the production efficiency at the laborer level (V^L) constant but increase the production efficiency at the CEO level (V^E), the total number of young workers hired increases. The intuition is that as the CEO job becomes more efficient, firms can extract more rents from the CEO worker if this worker is more able. Firms achieve this sorting by increasing the contestants' pool. I summarize this property in Proposition 6.

Proposition 6. *The total number of young workers hired in period 1 increases with V^E, holding V^L constant.*

From Proposition 6, a firm with a more efficient production technology at the CEO level hires more young workers and the market expectation about the CEO's ability increases with the number of young workers competing in the promotion tournament. Thus, if the CEO job becomes more efficient, the firm grows holding the firm structure constant. As shown in Proposition 5 , when firms grow, wages at all levels go up and the wage distribution becomes more unequal. I summarize these relationships in Proposition 7.

Proposition 7. *Consider two firms with two layers. Suppose* $V_1^E > V_2^E, V_1^L = V_2^L$*, then* $n_1 > n_2, W_1^j > W_2^j, j \in \{E, L\}$*, and* $\Delta W_1 > \Delta W_2$*.*

The wage patterns described in Proposition 7 are driven by the difference in the production technology. The two-period model shows that firms will adjust their sizes in response to technology changes while firm structure is held constant. That is, firms with more efficient production technology at the CEO level are larger. However, Rajan and Wulf (2006) find that firm sizes are relatively stable over time in their dataset although there has been substantial technology advancements during their sample period. This suggests that firms adjust to technology changes through changes in firm structure while holding firm size constant. Before I discuss how firms' structure changes with technology, let me first consider a three-layer model in the following section.

3.4.2 A three-layer model

In a three-layer firm, the CEO (E) occupies the top level, managers (M) stay on the middle management level, and laborers (L) take up the lower level. Each manager heads a division with a team of laborers. All managers report directly to the CEO. The production efficiency at each job level is denoted by $\hat{V}^j, j \in \{E, M, L\}$. Similar to the two layer case, an upper level job is more efficient in utilizing workers' ability than a lower level job, i.e., $\hat{V}^E > \hat{V}^M > \hat{V}^L$. The production function is the same as before. An individual with ability θ who is assigned to job j produces $S\hat{V}^j\theta$ units of output if she is old and works for her incumbent firm; $\hat{V}^j\theta$ otherwise.

Firms live for two periods. In period 1, $\hat{m}$ young workers are hired into the management positions in each firm and $\hat{n}$ young workers are hired into the laborer positions in each management division, i.e., $\hat{m} \cdot (\hat{n}+1)$ young workers are hired in each firm in period 1. In period 2, when the workers are old, one of the managers will be chosen to fill the CEO position. After the tournament at the management level, the laborers in the promoted manager's division compete for the management vacancy. I assume that the CEO job requires some manager-specific human capital that only those who have worked as a manager can work as a CEO. I also assume that there is some division-specific human capital that only the laborers in a particular division can work as the manager in that division.[7]

The firm's problem at the management level in the three-layer model is very similar to the one in the two-layer model. In period 2, after observing a manager's job assignment decisions made by her incumbent firm, outside firms form expectations about the manager's ability. I assume that the management production efficiency is not very different from the laborer production efficiency so that there are no demotions.[8] Thus, a manager is either promoted to be the CEO or remains on the management level. Denote the expected ability of the promoted manager (i.e., the CEO) $\theta_P^E(\hat{m})$ and the expected ability of a non-promoted manager $\theta_N^M(\hat{m})$. Then $\theta_P^E(\hat{m}) = E(\theta_i | i \text{ is the best among } \hat{m} \text{ managers})$ and $\theta_N^M(\hat{m}) = E(\theta_i | i$ is not the best among $\hat{m}$ managers$)$.

At the laborer level, only one laborer in the division with a management vacancy will be promoted to become a manager. Outside firms only observe work-

[7]This assumption rules out cross-division promotions. Otherwise there are more laborers competing for a single management position than managers competing for the CEO position such that the wage of the winning laborer exceeds the wage of the CEO, which is counterfactual. See Friebel and Raith (2013) for a theoretical analysis where cross-division promotions are allowed.

[8]It is possible that firms would want to demote a manager and replace her with a more competent laborer after their true abilities are revealed. However, in doing so, the incumbent firm would send a strong signal to the market about the laborer's ability such that the productivity gain could not compensate for the increase in the wage bill. Thus, if the manager-level production efficiency is not sufficiently different from the laborer-level production efficiency, demotions will not occur.

ers' job assignments but they do not observe in which division a management vacancy becomes available. Since all firms are identical, firms know the number of equilibrium young managers that other firms hire. Thus, if a laborer is promoted to be a manager, the expected ability of the promoted laborer is $\frac{1}{\hat{m}}\theta_P^M(\hat{n}) = \frac{1}{\hat{m}}E(\theta_i|i \text{ is the best among } \hat{n} \text{ laborers })$. The expected ability of the non-promoted laborer is $\frac{1}{\hat{m}}\theta_N^L(\hat{n}) + (1-\frac{1}{\hat{m}})E(\theta) = \frac{1}{\hat{m}}E(\theta_i|i \text{ is not the best among } \hat{n} \text{ laborers}) + (1-\frac{1}{\hat{m}})E(\theta)$.

Now let us consider the second-period wages in a three-layer model. $\hat{W}_P^E(\hat{V},\hat{m})$ is the wage for a CEO who is promoted from a management position. $\hat{W}_N^M(\hat{V},\hat{m})$ is the wage for a manager who is not promoted to be a CEO from a management position. $\hat{W}_P^M(\hat{V},\hat{m},\hat{n})$ is the wage for a manager who is just promoted to be a manger from a laborer position. $\hat{W}_N^L(\hat{V},\hat{m},\hat{n})$ is the wage for a laborer who is not promoted to be a manger from a laborer position. Proposition 8 describes the wages for the old workers in three-layer firms.

Proposition 8. *In equilibrium, an old worker's wage is equal to her expected productivity at an outside firm given her job assignment. At the management level, a promoted manager's wage is* $\hat{W}_P^E(\hat{V},\hat{m}) = \hat{V}^L\theta_P^E(\hat{m})$ *and a non-promoted manager's wage is* $\hat{W}_N^M(\hat{V},\hat{m}) = \hat{V}^L\theta_N^M(\hat{m})$. *At the laborer level, a promoted laborer's wage is* $\hat{W}_P^M(\hat{V},\hat{m},\hat{n}) = \hat{V}^L[\frac{1}{\hat{m}}\theta_P^M(\hat{n})]$; *a non-promoted laborer's wage is* $\hat{W}_N^L(\hat{V},\hat{m},\hat{n}) = \hat{V}^L[\frac{1}{\hat{m}}\theta_N^L(\hat{n})+(1-\frac{1}{\hat{m}})E(\theta)]$.

I now consider the first-period problem in each firm. In period 1, at the management level, firms choose a wage for young managers and decide how many managers to hire into the management position. Let $\hat{m}$ denotes the number of young managers to hire. $\hat{W}_Y^M$ denote the first period's wage for young managers, which is a function of $\hat{m}$. The problem at the management level is the following.

$$\underset{\hat{m}}{Max} \quad \hat{m}\hat{V}^M E(\theta) - \hat{m}\hat{W}_Y^M(\hat{m}) + [S\hat{V}^E \theta_P^E(\hat{m}) - \hat{W}_P^E(\hat{V},\hat{m})] \tag{3.4}$$

$$+(\hat{m}-1)[S\hat{V}^M \theta_N^M(\hat{m}) - \hat{W}_N^M(\hat{V},\hat{m})]$$

$$s.t. \quad \hat{W}_Y^M(\hat{m}) + [\tfrac{1}{\hat{m}}\hat{W}_P^E(\hat{V},\hat{m}) + \tfrac{\hat{m}-1}{\hat{m}}\hat{W}_N^M(\hat{V},\hat{m})] \geq \bar{U}_1 + \bar{U}_2 \tag{3.5}$$

(3.5) is a worker's participation constraint. $\bar{U}_1 + \bar{U}_2$ is a worker's expected lifetime income if she stays self-employed. That is, the young managers' wage is such that she is indifferent between working or staying self-employed for two periods. The first-order condition that characterizes the problem in (3.4) is

$$S(\hat{V}^E - \hat{V}^M)\frac{\theta_H - \theta_L}{(\hat{m}+1)^2} = (\bar{U}_1 + \bar{U}_2) - (1+S)\hat{V}^M E(\theta). \tag{3.6}$$

Note that the total number of young managers to hire is independent of the number of laborers and is determined only by the production efficiencies at the CEO level and the management level.

The problem at the laborers' level is similar to the problem at the management level. The only difference is that there might be no vacancy in a laborer's division. Let $\hat{n}$ denote the number of young laborers to hire. $\hat{W}_Y^L(\hat{n})$ denotes the first period's wage for young laborers. The problem at the laborer level is the following.

$$\underset{\hat{n}}{Max} \quad \hat{n}\hat{V}^L E(\theta) - \hat{n}\hat{W}_Y^L(\hat{n}) + (1-\frac{1}{\hat{m}})[\hat{n}S\hat{V}^L E(\theta) - \hat{n}\hat{W}_N^L(\hat{V},\hat{m},\hat{n})] \tag{3.7}$$

$$+\frac{1}{\hat{m}}\{[S\hat{V}\theta_P^M(\hat{n}) - \hat{W}_P^M(\hat{V},\hat{m},\hat{n})] + (\hat{n}-1)[S\hat{V}^L\theta_N^L(\hat{n}) - \hat{W}_N^L(\hat{V},\hat{m},\hat{n})]\}$$

$$s.t. \quad \hat{W}_Y^L(\hat{n}) + (1-\frac{1}{\hat{m}})\hat{W}_N^L(\hat{V},\hat{m},\hat{n})$$

$$+\frac{1}{\hat{m}}[\frac{1}{\hat{n}}\hat{W}_P^M(\hat{V},\hat{m},\hat{n}) + \frac{\hat{n}-1}{\hat{n}}\hat{W}_N^L(\hat{V},\hat{m},\hat{n})] \geq \bar{U}_1 + \bar{U}_2 \tag{3.8}$$

Similar to the manager's problem, (3.8) is a worker's participation constraint. The first-order condition that characterizes the problem in (3.7) is

$$\frac{1}{\hat{m}} S(\hat{V}^M - \hat{V}^L) \frac{\theta_H - \theta_L}{(\hat{n}+1)^2} = (\bar{U}_1 + \bar{U}_2) - (1+S)\hat{V}^L E(\theta). \tag{3.9}$$

From (3.9), the more young managers there are, the fewer young laborers there are in each division. The intuition is that if there are many managers, the probability that a vacancy occurs is low. Thus, it is less likely that a firm can extract the rent by placing a more able worker into the management position. As a result, firms hire fewer laborers.

Compare the period-1's problem at the management level and at the laborer level. If the difference in terms of production efficiency between the CEO job and the manager job is larger than the difference between the manager job and the laborer job, the CEO would have a larger span-of-control than the managers. That is,

Proposition 9. *If* $\hat{V}^E - \hat{V}^M > \hat{V}^M - \hat{V}^L$, *then* $\hat{m} > \hat{n}$.

Proposition 9 says that upper-level managers have a larger span than lower-level managers. This is due to the mechanism that firms have the incentive to assign more able workers to more efficient jobs and the sorting is achieved through increasing the number of contestant in a promotion tournament.

3.4.3 Technology, delayering, and wages

Now, we are ready to analyze the relationship between delayering and wages. Let us first look at the relationship between technology changes, firm sizes, and firm structure. We know from the delayering literature that one of the main causes of firm delayering is technology advancement. Thus, I consider an exogenous technology shock that changes the production efficiency at the CEO level in a three-layer firm from $\hat{V}^E$ to V^E and the firm decides to restructure into a two-layer firm. Since Rajan and Wulf (2006) find that firms' sizes are relatively stable over the years, I assume that the firm now hires the same number of workers as before, i.e., $N = \hat{N}$. I have the following Proposition.

Proposition 10. *If firm size and the production efficiency at the laborer level are unchanged, the CEO's production efficiency must be higher after delayering, i.e. if* $\hat{N} = N, \hat{V}^L = V^L$*, then it must be that* $V^E > \hat{V}^E$.

The logic behind Proposition 10 is the following. If a firm only has two layers, all workers are led by the CEO. If a firm has three layers, the CEO only leads the managers. If the firm has the same number of workers after delayering, it must be that the CEO in the two-layer firm leads a larger production team, which means she is more efficient. Note that Proposition 10 considers within firm changes rather than a cross firm comparison.

Now, let us consider delayering and wage changes. If we compare the non-promoted laborers' wage, $\hat{W}^L_N$, in the three-layer model to that in the two-layer model, W^L, we can see that the laborers' wage is higher in a two-layer firm, holding firm size and the production efficiency at the laborer level constant. The reason is that the two-layer firm has a larger contestant pool such that the expected ability of a loser is higher than the expected ability of a loser in a smaller contest. Since the production efficiency at the laborer level is unchanged, the flatter firm pays a higher non-promotion wage to the laborers. Similar argument applies to the wages paid to the CEOs in the two-layer firm and the three-layer firm. That is, after delayering, the CEO and the laborers are paid more because of an upward adjustment in their expected abilities.

With regard to the wage difference between the CEO and the laborer, when the market expectation about the non-promoted laborers' ability increases, the market expectation about the CEO's ability also increases. With a uniform ability distribution, the difference between these two expectations becomes larger as the number of contestants gets larger. That is, after delayering, the wage distribution within a firm becomes more unequal. I summarize the above argument in Proposition 11.

Proposition 11. *Suppose there is a technology change that results in a higher CEO production efficiency, i.e.,* $V^E > \hat{V}^E$*, and delayering. If firm size and the production efficiency at the laborer level are unchanged, both the laborers and the CEO's wages increase, i.e., if* $\hat{N} = N, \hat{V}^L = V^L$*, then* $W^E > \hat{W}^E_P, W^L > \hat{W}^L_N$*. In addition,* $\Delta W = W^E - W^L > \Delta \hat{W} = \hat{W}^E_P - \hat{W}^L_N$

The predictions in Proposition 11 capture the empirical findings in the delayering literature that when firms become flatter, wages at all level go up (Bauer and Bender, 2001; Rajan and Wulf, 2006; Caliendo et al., 2012) and wage inequality increases (Bauer and Bender, 2001).

In Zabojnik and Bernhardt (2001), more efficient firms are larger given that all firms have the same number of layers. From Proposition 10, a more efficient firm (measured by the production efficiency at the CEO level) is not necessarily larger than a less-efficient firm if we take into account firm restructuring. This result is consistent with the empirical finding in Rajan and Wulf (2006) that firms have fewer layers without much changes in firm size over time.

3.5 Conclusion

This paper develops a model with asymmetric learning and slot constraints to explore the relationship between firm delayering and workers' wages. It contributes to the literature in multiple ways. First, it contributes to the delayering literature by exploring the consequences of delayering on wages while most of the delayering literature focuses on the causes of the delayering trend. Second, it contributes to the job assignment literature by considering how firms' organization structure affects wages. Third, this model captures several empirical findings that are not well explained in the existing literature. My model shows that after delayering, workers' wages at all levels increase because they are now participating in a larger contest and thus their expected abilities are higher. In addition, since the workers' wage at the top increases faster than the wage increase at the bottom, the wage distribution becomes more unequal after delayering.

APPENDIX A

APPENDIX FOR CHAPTER 1

Derivation of the parameter restriction in equation (1.3).

Let us consider the four possible equilibrium career development paths (all other paths are strictly dominated as explained in the text) for a type-j worker and her total productivity in all three periods associated with each of the four career paths. Since in each path, the productivity in period 1 is the same, we simply consider the total productivity in period 2 and 3. Since there is free entry, firms make zero profit by paying out all the expected output as wages. Thus, firms' problem is to maximize expected productivity by assigning workers optimally. $\Pi_{j,2,2}$ denote the total productivity (in period 2 and 3) of a type-j worker who starts in job j, is promoted to level 2 in period 2, and stays on level 2 in period 3. $\Pi_{j,j,2}$ denote the total productivity of a type-j worker who starts in job j, stays in job j in period 2, and is promoted to level 2 in period 3. This is the productivity of the non-movers. $\Pi_{j,k,2}$ denote the total productivity of a type-j worker who starts in job j, is laterally moved to job k in period 2, and is promoted to level 2 in period 3. This is the productivity of the movers. Finally, $\Pi_{j,j,j}$ denote the total productivity of a type-j worker who starts in job j and stays in job j for all three periods. We have the following expressions.

$$\Pi_{j,2,2}(\theta) = \{d_2 + c_2\theta[f(\gamma(1+\alpha)+s) + f(s)]\} + \{d_2 + c_2\theta[f(1+\gamma(1+\alpha)+s) + f(1+s)]\} \tag{A.1}$$

$$\Pi_{j,j,2}(\theta) = \{d_1 + c_1\theta f(1+\alpha+s)\} + \{d_2 + c_2\theta[f(\gamma(2+\alpha)+s) + f(s)]\} \tag{A.2}$$

$$\Pi_{j,k,2}(\theta) = \{d_1 + c_1\theta f(s)\} + \{d_2 + c_2\theta[f(\gamma(1+\alpha)+s) + f(\gamma+s)]\} \tag{A.3}$$

$$\Pi_{j,j,j}(\theta) = \{d_1 + c_1\theta f(1+\alpha+s)\} + \{d_1 + c_1\theta f(2+\alpha+s)\} \tag{A.4}$$

We can see that for any positive θ, either $\Pi_{j,j,2} > \Pi_{j,k,2}$ or $\Pi_{j,k,2} > \Pi_{j,j,2}$. Parameters in (1.3) guarantees that the latter inequality holds that lateral moves exist in equilibrium. $\|$

A.1 Proof of Proposition 1

Equating (A.1) and (A.3) gives the cut off ability to promote a worker or not in period 2. Let θ^* satisfy $\Pi_{j,2,2}(\theta^*) = \Pi_{j,k,2}(\theta^*)$. Solve for θ^*, we get

$$\theta^* = \frac{d_1 - d_2}{c_2[f(1+\gamma(1+\alpha)+s) + f(1+s) + f(s) - f(\gamma+s)] - c_1 f(s)}.$$

For $\theta_i > \theta^*$, $\Pi_{j,2,2} > \Pi_{j,k,2}$. Thus, it is optimal to promote a worker in period 2. Similarly, equating (A.3) and (A.4) gives the cut off ability to lateral move a worker or not in period 2. Let $\tilde{\theta}^*$ satisfy $\Pi_{j,k,2}(\tilde{\theta}^*) = \Pi_{j,j,j}(\tilde{\theta}^*)$. Solve for $\tilde{\theta}^*$, we get

$$\tilde{\theta}^* = \frac{d_1 - d_2}{c_2[f(\gamma(1+\alpha)+s) + f(\gamma+s)] - c_1[f(2+\alpha+s) + f(1+\alpha+s) - f(s)]}.$$

For $\theta_i > \tilde{\theta}^*$, $\Pi_{j,k,2} > \Pi_{j,j,j}$. Thus, it is optimal to move a worker in period 2 and promote her in period 3. $\|$

A.2 Proof of Proposition 2

In period 3, firms problem is to maximize the last period productivity given workers' job assignment history. There are two types of workers with different career paths that are considered promotion. Let θ_3^{jk} and θ_3^{jj} denote, respectively, the cutoff ability levels for promotions in period 3 for a period-2 mover and a period-2 non-mover (a non-mover is not promoted or moved). θ_3^{jk} solves $d_1 + c_1\theta_3^{jk} f(1+\alpha+s) = d_2 + c_2\theta_3^{jk}[f(\gamma+\gamma\alpha+$

$s)+f(\gamma+1+s)]$, which says θ_3^{jk} is such that firms are indifferent between promoting a period-2 mover or let her move back to her matched job. For $\theta_{i3}^e > \theta_3^{jk}$, a period-2 mover is more productive if promoted to level 2. Otherwise, she is more productive in her matched job. $\theta_3^{jk} = (d_1-d_2)/\{c_2[f(s+\gamma+\gamma\alpha)+f(s+\gamma)]-c_1 f(1+\alpha+s)\}$. Similarly, θ_3^{jj} solves $d_1+c_1\theta_3^{jj}f(2+\alpha+s) = d_2+c_2\theta_3^{jj}[f(2\gamma+\gamma\alpha+s)+f(s)]$, which says θ_3^{jj} is such that firms are indifferent between promoting a period-2 non-mover or make her stay in her matched job. $\theta_3^{jj} = (d_1-d_2)/\{c_2[f(s+2\gamma+\gamma\alpha)+f(s)]-c_1 f(2+\alpha+s)\}$. We can see that, $\theta_3^{jj} > \theta_3^{jk}$. This proves (iv). Let θ_3^{j2} denote the cutoff ability level for a period-2 promoted worker to stay on level 2 in period 3. θ_3^{j2} solves $d_1+c_1\theta_3^{j2}f(1+\alpha+s) = d_2+c_2\theta_3^{j2}[f(1+\gamma+\gamma\alpha+s)+f(1+s)]$, $\theta_3^{j2} = (d_1-d_2)/\{c_2[f(1+s+\gamma+\gamma\alpha)+f(1+s)]-c_1 f(1+\alpha+s)\}$. If $\theta_{i3}^e > \theta_2^{j2}$, the worker remains on level 2, otherwise she is demoted into her matched job on level 1. This proves (i)

In period 2, θ_2^j denotes the cutoff ability level for promotion, which solves

$$
\begin{aligned}
&\{d_1+c_1\theta_2^j f(s)\}+prob.(\theta_{i3}^e \geq \theta_{i3}^{jk}|\theta_{i2}^e=\theta_2^j)\{d_2+c_2\theta_2^j[f(s+\gamma+\gamma\alpha)+f(\gamma+s)]\} \\
&\qquad +prob.(\theta_{i3}^e < \theta_{i3}^{jk}|\theta_{i2}^e=\theta_2^j)\{d_1+c_1\theta_2^j f(1+\alpha+s)\} \\
&= \{d_2+c_2\theta_2^j[f(s+\gamma+\gamma\alpha)+f(s)]\}+prob.(\theta_{i3}^e \geq \theta_3^{j2}|\theta_{i2}^e=\theta_2^j)\{d_2+c_2\theta_2^j[f(1+s+\gamma+\gamma\alpha) \\
&\qquad +f(1+s)]\}+prob.(\theta_{i3}^e < \theta_3^{j2}|\theta_{i2}^e=\theta_2^j)\{d_1+c_1\theta_2^j f(1+s+\alpha)\}.
\end{aligned}
$$

That is, θ_2^j is such that firms are indifferent between laterally moving a worker or promoting her in period 2. For $\theta_{i2}^e > \theta_2^j$, it is more productive to promote a worker. This proves (ii). (iii) is straightforward from the parameterization. ||

A.3 Proof of Corollary 1

$X = prob.(\theta_{i3}^e > \theta_3^{jk} | \tilde{\theta}_2^j < \theta_{i2}^e \leq \theta_2^j)$ denote the probability that a laterally moved worker is promoted in period 3 from period 2's perspective. $Y = prob.(\theta_{i3}^e > \theta_3^{jj} | \theta_{i2}^e < \tilde{\theta}_2^j)$ is the probability that a non-mover is promoted in period 3. At the beginning of period 1, $P_1 := prob.(\theta = \theta_H) = p_0$. $z_{i1} = \theta_i f(s+\alpha) + \varepsilon_{ij1} = (y_{ij1} - d_1)/c_1$. Define

$$\begin{aligned} P_2 := prob.(\theta = \theta_H | z_{i1}) &= \frac{p_0 h[z_{i1} - \theta_H f(s+\alpha)]}{p_0 h[z_{i1} - \theta_H f(s+\alpha)] + (1-p_0) h[z_{i1} - \theta_L f(s+\alpha)]} \\ &= \frac{p_0}{p_0 + (1-p_0)\frac{h[z_{i1} - \theta_L f(s+\alpha)]}{h[z_{i1} - \theta_H f(s+\alpha)]}} \end{aligned}$$

, where $h(\cdot)$ is the density of the error term. Then, given P_2, z_{i2} is such that

$$\begin{aligned} P_3 := prob.(\theta = \theta_H | z_{i1}, z_{i2}) &= \frac{P_2 h[z_{i2} - \theta_H f(s)]}{P_2 h[z_{i2} - \theta_H f(s)] + (1-P_2) h[z_{i2} - \theta_L f(s)]} \\ &= \frac{P_2}{P_2 + (1-P_2)\frac{h[z_{i2} - \theta_L f(s)]}{h[z_{i2} - \theta_H f(s)]}}. \end{aligned}$$

Since $\frac{h[z_{i1} - \theta_L f(\cdot)]}{h[z_{i1} - \theta_H f(\cdot)]}$ decreases in z_{i2}, P_3 increases in z_{i2}, there exists a cutoff probability P_3^{lat*} such that when $P_3 > P_3^{lat*}$, a period-2 mover is promoted in period 3. P_3^{lat*} satisfies $P_3^{lat*}\theta_H + (1 - P_3^{lat*})\theta_L = \theta_3^{jk}$, i.e., $P_3^{lat*} = (\theta_3^{jk} - \theta_L)/(\theta_H - \theta_L)$. That is, given P_2, there exists a $z_3^{lat*}(P_2)$ such that if $z_{i2} > z_3^{lat*}(P_2)$, a period-2 mover is promoted in period 3. Similary, there exists a cutoff probability P_3^{stay*} such that when $P_3 > P_3^{stay*}$, a period-2 non-mover is promoted in period 3. P_3^{stay*} satisfies $P_3^{stay*}\theta_H + (1 - P_3^{stay*})\theta_L = \theta_3^{jj}$, i.e., $P_3^{stay*} = (\theta_3^{jj} - \theta_L)/(\theta_H - \theta_L)$. Then, there exisits a $z_3^{stay*}(P_2)$ such that if $z_{i2} > z_3^{stay*}(P_2)$, a period-2 non-mover is promoted in period 3. As shown in Proposition 2 (iv), $\theta_3^{jk} < \theta_3^{jj}$, thus $P_3^{lat*} < P_3^{stay*}$.

Furthermore, $prob.(z_{i2} \geq z(p)), \forall z(p)$, increases in p, since lateral move means a more positive P_2 and the promotion cutoff for a period-2 mover is lower, the probability of promoting a mover is higher. ‖

A.4 Proof of Corollary 2

Consider the wage changes from period 2 to period 3 after the lateral move. Let w_{Lt} denote the wage in t for a period-2 mover. w_{St} denote the wage in t for a period-2 non-mover. $\theta^e_{i1} = \theta_0, \forall i; \theta^e_{i2} = E(\theta|z_{i1}), \theta^e_{i3} = E(\theta|z_{i1}, z_{i2}), i \in \{S, L\}$. In period 2, the expected wage change from period 2 to period 3 for a mover (L) is

$$\begin{aligned}\Delta w^L := E(w_{L3} - w_{L2}|P_1, P_2) &= c_1\theta^e_{L2}[f(1+\alpha+s) - f(s)] \\ &+ E[g(\theta^e_{L3})|P_1, P_2]\{c_2[f(\gamma+\gamma\alpha+s) + f(\gamma+s)] \\ &- c_1 f(1+\alpha+s)\},\end{aligned}$$

where $g(\theta^e_{L3}) = max\{0, \theta^e_{L3} - \theta^{jk}_3\}$. The expected wage change from period 2 to period 3 for a non-mover (S) is

$$\begin{aligned}\Delta w^S := E(w_{S3} - w_{S2}|P_1, P_2) &= c_1\theta^e_{S2}[f(2+\alpha+s) - f(1+\alpha+s)] \\ &+ E[g(\theta^e_{S3})|P_1, P_2]\{c_2[f(2\gamma+\gamma\alpha+s) + f(s)] \\ &- c_1 f(2+\alpha+s)\}\end{aligned}$$

, where $g(\theta^e_{S3}) = max\{0, \theta^e_{S3} - \theta^{jj}_3\}$. Since $\forall\theta^*, prob.(\theta^e_{i3} \geq \theta^*|P_1, P_2)$ increases in P_2 and $\theta^{jk}_3 < \theta^{jj}_3$, we have $E[g(\theta^e_{L3})|P_1, P_2] > E[g(\theta^e_{S3})|P_1, P_2]$. Using the concavity of $f(\cdot)$ and the fact that $\theta^e_{L2} > \theta^e_{S2}$, we have $\Delta w^L > \Delta w^S$.

Now let's consider briefly the wage changes from period 1 to period 2 in the period of lateral move. In period 1, the expected wage change from period 1 to period 2 for a mover is $E(w_{L2} - w_{L1}|P_1) = c_1[\theta^e_{L2} f(s) - \theta_0 f(s+\alpha)]$; the expected wage change from period 1 to period 2 for a non-mover is $E(w_{S2} - w_{S1}|P_1) = c_1[\theta^e_{S2} f(1+s+\alpha) - \theta_0 f(s+\alpha)]$. Since $\theta^e_{L2} > \theta^e_{S2}$, the wage relation is ambiguous. $\|$

Table A.1: Lateral Moves by Years and Reporting Level

	Lateral Move (t)					
Level (t)	*t = 1982*		*t = 1983*		*t = 1984*	
	Freq.	Pct.	Freq.	Pct.	Freq.	Pct.
1	4	2.0%	8	3.3%	5	2.0%
2	101	10.2%	123	10.0%	114	9.1%
3	262	9.7%	406	11.9%	360	10.2%
4	439	12.1%	706	14.8%	693	14.1%
5	308	12.1%	482	13.6%	498	14.3%
6	115	10.6%	194	12.5%	231	15.8%
7	28	7.3%	53	10.2%	77	15.7%
8	5	4.8%	10	6.1%	3	1.9%
9	0	0.0%	3	4.8%	0	0.0%
10	0	0.0%	1	4.6%	0	0.0%
11	0	0.0%	0	0.0%	0	0.0%
Total	1,262	10.8%	1,986	12.8%	1,981	12.7%

A.5 Robustness Checks

In the main analysis, I use past-period's total pay as a proxy for individuals' expected productivity. In Table **??**, I present the results replicating Columns (3) and (6) in Table 1.5. The effect of lateral moves on promotions is robust to different measures of expected productivity. In particular, two years after the move, the movers are more likely to earn a promotion.

In Table A.3, I present the results replicating Columns (3) and (6) in Table 1.5 clustering the standard errors by individual and by firm using Bootstrap. The lateral-more-promotion relation persists.

In Table A.4, I present the results replicating Table 1.6 by clustering the standard errors at the firm level. The lateral-move-wage relation persists.

Table A.2: Lateral Moves and Promotion Probability: Different Measures of Expected Productivity

Dependent Variable:	Lateral Move in 82 (τ=82;t=83,84,85)			Lateral Moves in 83 (τ=83;t=84,85)		
Promotion (t)=1 if Yes	(1)	(2)	(3)	(4)	(5)	(6)
Exp. Prod. Measure (10k)	Base	Midgrade	Mid./ Age	Base	Midgrade	Mid./ Age
Lateral move effect by year						
Lateral (τ)*Year1	-0.015*	-0.008	-0.009	0.008	0.004	0.003
	(0.009)	(0.011)	(0.011)	(0.010)	(0.012)	(0.011)
Lateral (τ)*Year2	0.031**	0.036**	0.035**	0.035**	0.039**	0.039**
	(0.013)	(0.015)	(0.015)	(0.015)	(0.018)	(0.018)
Latearl (τ)*Year3	0.029*	0.028	0.029	-	-	-
	(0.016)	(0.019)	(0.019)	-	-	-
Past in-sample observables						
Exp. Productivity (t-1)	0.010***	0.005***	0.275***	0.014***	0.012***	0.554***
	(0.002)	(0.002)	(0.099)	(0.002)	(0.002)	(0.114)
Levels (t-1)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Unemp. Rate (t-1)	0.008	0.009	0.009	-0.004	-0.009	-0.008
	(0.008)	(0.009)	(0.009)	(0.011)	(0.015)	(0.014)
Firm Char. (t-1)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Individual Initial Char.						
Education	0.004***	0.005***	0.004***	0.006***	0.006***	0.005***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age	-0.002***	-0.001***	0.001**	-0.003***	-0.003***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Job Tenure	-0.002**	-0.001*	-0.001*	-0.003***	-0.002*	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Employer Tenure	-0.001	-0.001**	-0.001*	-0.001**	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unemp Rate	-0.009	-0.010	-0.010	-0.000	0.003	0.003
(at hiring)	(0.008)	(0.009)	(0.009)	(0.011)	(0.015)	(0.014)
Exp. Productivity	0.003	0.009***	0.483***	0.002	0.005**	0.314***
(initial)	(0.002)	(0.002)	(0.102)	(0.002)	(0.002)	(0.113)
Functional Areas	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
σ_u	0.00494	0.00372	0.00330	0.544	0.557	0.509
ρ	2.44e-05	1.38e-05	1.09e-05	0.228	0.237	0.206
N. of Observations	19,261	15,899	15,899	13,451	11,073	11,073
N. of firm-individuals	9,586	8,010	8,010	8,644	7,150	7,150
N. of quadrature points	12	12	12	12	12	12
Log likelihood	-6575	-5645	-5639	-4593	-3946	-3948

Note 1- Standard errors are in parentheses. Average Marginal Effects are reported. Marginal effects for dummies are the differences in predicted probabilities when the dummy equals 1 and when it equals 0.
Note 2 - Sample restricted to executives who are not promoted in the current period.
* Statistically significant at the 10% level. ** Statistically significant at the 5% level.
*** Statistically significant at the 1% level.

Table A.3: Lateral Moves and Promotion Probability After the moves: Probit Models with Clustered Standard Errors

Dependent Variable:	Lateral Move in 82 (τ=82;t=83,84,85)				Lateral Moves in 83 (τ=83;t=84,85)			
Promotion (t)=1 if Yes	(1) cluster S.E. by ind.		(2) cluster S.E. by firm		(3) cluster S.E. by ind.		(4) cluster S.E. by firm	
	Coef.	Margins	Coef.	Margins	Coef.	Margins	Coef.	Margins
Lateral move effect by year								
Lateral (τ)*Year1	-0.264***	-0.034***	-0.264**	-0.034***	-0.061	-0.007	-0.061	-0.007
	(0.096)	(0.011)	(0.111)	(0.013)	(0.137)	(0.016)	(0.142)	(0.017)
Lateral (τ)*Year2	0.257***	0.043**	0.257**	0.043**	0.315**	0.046**	0.315*	0.046
	(0.092)	(0.017)	(0.114)	(0.021)	(0.143)	(0.023)	(0.189)	(0.034)
Latearl (τ)*Year3	0.412***	0.074***	0.412***	0.074**	-	-	-	-
	(0.113)	(0.023)	(0.142)	(0.030)	-	-	-	-
Past in-sample observables								
Total Pay (10k)(t-1)	0.034***	0.005***	0.034**	0.005**	0.063***	0.008***	0.063***	0.008***
	(0.010)	(0.001)	(0.014)	(0.002)	(0.021)	(0.002)	(0.018)	(0.002)
Levels (t-1)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Unemp. Rate (t-1)	0.058	0.009	0.058	0.009	-0.003	-0.000	-0.003	-0.000
	(0.069)	(0.010)	(0.097)	(0.014)	(0.135)	(0.017)	(0.137)	(0.017)
Firm Char. (t-1)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Individual Initial Char.								
Education	0.046***	0.007***	0.046***	0.007***	0.066***	0.008***	0.066***	0.008***
	(0.014)	(0.002)	(0.015)	(0.002)	(0.024)	(0.003)	(0.023)	(0.003)
Age	-0.014***	-0.002***	-0.014***	-0.002***	-0.016**	-0.002**	-0.016***	-0.002***
	(0.004)	(0.001)	(0.004)	(0.001)	(0.006)	(0.001)	(0.006)	(0.001)
Job Tenure	-0.012*	-0.002*	-0.012	-0.002	-0.021*	-0.003*	-0.021	-0.003
	(0.007)	(0.001)	(0.008)	(0.001)	(0.012)	(0.002)	(0.013)	(0.002)
Employer Tenure	-0.001	-0.000	-0.001	-0.000	-0.006	-0.001	-0.006	-0.001
	(0.004)	(0.001)	(0.004)	(0.001)	(0.006)	(0.001)	(0.005)	(0.001)
Unemp Rate	-0.068	-0.010	-0.068	-0.010	-0.028	-0.003	-0.028	-0.003
(at hiring)	(0.068)	(0.010)	(0.097)	(0.014)	(0.136)	(0.017)	(0.139)	(0.017)
Total Pay (10k)	0.029***	0.004***	0.029*	0.004*	0.019	0.002	0.019	0.002
(initial)	(0.011)	(0.002)	(0.016)	(0.002)	(0.020)	(0.003)	(0.019)	(0.002)
Functional Areas	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Constant	-0.528*		-0.528*		-0.640		-0.293	
	(0.311)		(0.317)		(0.535)		(1.451)	
σ_u	0.894		0.894		1.138		1.138	
ρ	0.444		0.444		0.564		0.564	
N. of Observations	19,261		19,261		13,451		13,451	
N. of firm-individuals-year	4,666		4,666		4,669		4,669	
N. of quadrature points	12		12		12		12	
Log likelihood	-6254		-6254		-4340		-4601	

Note 1 - Robust standard errors are calculated using Bootstrap. Average Marginal Effects are reported. Marginal effects for dummies are the differences in predicted probabilities when the dummy equals 1 and when it equals 0.
Note 2 - Sample restricted to executives who are not promoted in the current period.
* Statistically significant at the 10% level. ** Statistically significant at the 5% level. *** Statistically significant at the 1% level.

Table A.4: Lateral Moves and Real Compensation Changes After the Moves (in 1982 dollar): RE Model with Clustered S.E. by Firm

	Lateral Moves in 82 (τ=82; t=83,84,85)			Lateral Moves in 83 (τ=83; t=84,85)		
Real Comp. Change from the Year of the Moves:	(1) Total (t-τ)	(2) Base (t-τ)	(3) Bonus (t-τ)	(4) Total (t-τ)	(5) Base (t-τ)	(6) Bonus (t-τ)
Lateral move effect by year						
Lateral (τ)*Year1	-2,202.934**	-1,866.735***	-320.140	492.100	-836.602**	1,359.234*
	(996.547)	(387.435)	(865.267)	(812.230)	(376.499)	(707.299)
Lateral (τ)*Year2	4,785.421***	1,989.888***	2,769.327**	6,198.885***	2,796.888***	3,434.932***
	(1,380.115)	(554.903)	(1,212.391)	(1,327.990)	(583.884)	(1,228.157)
Latearl (τ)*Year3	10,530.765***	5,591.137***	5,006.637***	-	-	-
	(2,080.090)	(910.563)	(1,674.502)	-	-	-
Levels (t)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Δ Firm Char. (t)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Δ Unemp. Rate (t)	-3,306.002*	-1,208.606	-1,921.926**	-1,388.081	279.164	-1,517.386
	(1,978.257)	(1,131.027)	(906.875)	(1,845.631)	(1,001.465)	(1,045.782)
Individual Initial Char.						
Education	678.526***	321.536***	363.996**	679.296***	84.290	592.853***
	(203.365)	(88.985)	(170.465)	(183.627)	(79.957)	(167.826)
Age	184.906	395.443**	-204.658	387.254	67.554	326.922
	(360.993)	(161.653)	(278.874)	(292.378)	(138.755)	(223.257)
Age^2	-3.977	-6.046***	2.017	-6.404*	-2.325	-4.166
	(3.775)	(1.694)	(2.991)	(3.323)	(1.519)	(2.652)
Job tenure	-391.863***	-86.170	-303.887***	-296.599***	-129.549***	-165.858**
	(101.135)	(53.875)	(82.098)	(83.690)	(37.007)	(68.460)
Emp. tenure	115.003**	19.283	98.495**	109.303*	38.480**	71.874
	(57.644)	(24.866)	(48.087)	(57.160)	(17.779)	(50.405)
Total Pay (10k)	-206.397	204.052*	-429.507*	694.777***	320.929***	373.940**
(initial)	(284.662)	(120.577)	(225.675)	(185.244)	(56.356)	(177.605)
Functional Areas	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Constant	-8,032.538	-9,736.175**	1,360.725	-15,504.109*	807.118	-16,507.244**
	(8,645.608)	(4,322.859)	(6,315.330)	(7,949.427)	(3,692.365)	(6,497.921)
σ_u	19289	9781	14377	14452	6982	11921
σ_e	20113	7782	16667	16862	6108	14766
ρ	0.479	0.612	0.427	0.423	0.566	0.395
N. of Observations	19,416	19,416	19,416	13,512	13,512	13,512
N. of Clusters	9,586	9,586	9,586	8,644	8,644	8,644

Note 1 - Robust tandard errors are in parentheses; clustered by firm
Note 2 - Sample restricted to executives who are not promoted in the current period.
* Statistically significant at the 10% level.
** Statistically significant at the 5% level.
*** Statistically significant at the 1% level.

APPENDIX B

APPENDIX FOR CHAPTER 2

Expression for Parameterization (ii)

$$(1+S)\{E[y_{iT}^{m}|q_{iT}(1)]-E[y_{iT}^{l}|q_{iT}(1)]\} > E[y_{iT}^{m}|q_{iT}(1)]-E[y_{iT}^{l}|q_{iT}(0)].$$

This condition guarantees that a worker is promoted when a high productive efficiency is attained because the productivity gain from correct job assignment is larger than the loss in terms of paying a higher wage by sending the promotion signal to the market.

B.1 Proof of Proposition 3

In period T, the incumbent's and the outside firm's problem is to maximize the expected productivity in period T. Thus, a good worker is assigned to the upper-level job. In period T-1, in anticipating that a good worker is assigned to the upper-level job, the incumbent and the outside firms also maximize the expected productivity in period T-1 and assign a good worker to the upper-level job. By induction, a good worker is always assigned to the upper-level job. ||

B.2 Proof of Proposition 4

Since the incumbent's assignment decision is made by maximizing the current period productivity, given parameter restriction (i), all young workers are assigned to the lower level job. Now, observe that $q_t(0) < q_{t-1}(0) < \ldots < q_2(0)$, given parameter restriction (ii), if a worker continues to produce low, the expectation that she is good is below q^*. Therefore, a worker who produce low remains on level-l. Also, since the belief that a worker is good only depends on the total number of high productive efficiency that has been achieved up to a certain period, in period T, no matter when a worker attains a high output, the belief that she is good is the same. In addition, a worker with only one high output up to $T-1$ is less likely to be a good worker compared to workers with more high outputs. Thus, if a worker with only one high output in the previous $T-1$ periods is more productive on the upper level in expectation, all other types should be more productive on the upper level. Therefore, $q^* < q_T(1) < q_T(2) < \ldots < q_T(T-1)$ guarantees that once a worker attains high output, she is always more productive on the upper-level job. Parameter restriction (iii) guarantees that an incumbent firm gain more through correct job assignment and to assign a worker with at least one high output in the previous $T-1$ periods to the upper level job. $\|$

B.3 Proof of Corollary 3

From Proposition 4,

$$w_t^l[q_t(0)] = \{q_t(0)E_g^l + [1-q_t(0)]E_r^l + f(t-1)\}$$
$$+\beta S\{[q_t(0)\theta_g + (1-q_t(0))\theta_r][q_{t+1}(1)E_g^m + (1-q_{t+1}(1))E_r^m]$$
$$+[q_t(0)(1-\theta_g) + (1-q_t(0))(1-\theta_r)][q_{t+1}(0)E_g^l + (1-q_{t+1}(0))E_r^l] + f(t)\}$$

So $w_t^l[q_t(0)] - w_{t+1}^l[q_{t+1}(0)]$

$$= [q_t(0) - q_{t+1}(0)][E_g^l - E_r^l] - [f(t) - f(t-1)]$$
$$+\beta S\{E_t[Z_{t+1}^{J_{t+1}}|q_t(0),l] - E_{t+1}[Z_{t+2}^{J_{t+2}}|q_{t+1}(0),l] - [f(t+1) - f(t)]\}.$$

From Bayes' Rule, $q_t(0) = \frac{p_0(1-\theta_g)^{t-1}}{p_0(1-\theta_g)^{t-1}+(1-p_0)(1-\theta_r)^{t-1}}$. Thus,

$$\Delta q_t(0) = q_t(0) - q_{t+1}(0) = \frac{1}{1+\frac{1-p_0}{p_0}\left(\frac{1-\theta_r}{1-\theta_g}\right)^{t-1}} - \frac{1}{1+\frac{1-p_0}{p_0}\left(\frac{1-\theta_r}{1-\theta_g}\right)^{t}},$$

which decreases in t. Therefore, $q_2(0) - q_3(0) > q_t(0) - q_{t+1}(0)$. Similarly $q_t(1) - q_{t-1}(1) > q_{t+1}(1) - q_t(1), q_{t+1}(1) - q_t(1) > q_{t+1}(0) - q_t(0)$. Since $f(t) - f(t-1) > f(t_1^*) - f(t_1^* - 1) > f(t_1^*+1) - f(t_1^*) > [q_2(1) - q_3(1)][E_g^l - E_r^l] > [q_t(1) - q_{t+1}(1)][E_g^l - E_r^l] > [q_t(0) - q_{t+1}(0)][E_g^l - E_r^l]$ for $t < t_1^* + 1$, $w_t^l < w_{t+1}^l$. Since $f(t+1) - f(t) < f(t_1^* + 2) - f(t_1^* + 1) < [q_{\bar{t}}(1) - q_{\bar{t}+1}(1)][E_g^l - E_r^l] < [q_t(1) - q_{t+1}(1)][E_g^l - E_r^l]$ for $t_1^* + 1 < t \leq \bar{t}$ and $f(t+1) - f(t) \to 0$ for $t \geq \bar{t}$, $w_t^l > w_{t+1}^l$. $\|$

B.4 Proof of Corollary 4

For a promoted worker who attains H in t, in expectation, this worker can produce H or L in the next time but her expected productivity is fixed at the time t when she attains H. Thus,

$$\begin{aligned} w_t^m[q_t(1)] = &\{q_t(1)E_g^m + [1-q_t(1)]E_r^m + f(t-1)\} \\ &+ \beta\{(1+S)[q_t(1)\theta_g + (1-q_t(1))\theta_r][q_{t+1}(2)E_g^m + (1-q_{t+1}(2))E_r^m] \\ &+ (1+S)[q_t(1)(1-\theta_g) + (1-q_t(1))(1-\theta_r)][q_{t+1}(1)E_g^m + (1-q_{t+1}(1))E_r^m] \\ &- \{q_t(1)E_g^m + [1-q_t(1)]E_r^m\} + Sf(t)\} \end{aligned}$$

So $w_t^m[q_t(1)] - w_{t+1}^m[q_{t+1}(1)]$

$$\begin{aligned} &= (1-\beta)[q_t(1) - q_{t+1}(1)][E_g^m - E_r^m] - [f(t) - f(t-1)] \\ &+ \beta\{(1+S)\left[E_t[Z_{t+1}^{J_{t+1}}|q_t(1), m] - E_{t+1}[Z_{t+2}^{J_{t+2}}|q_{t+1}(1), m]\right] - S[f(t+1) - f(t)]\}. \end{aligned}$$

Similar to the proof for Corollary 4, $\Delta q_{t+1}(2) > \Delta q_{t+2}(2)$ for $t \geq 3$. Using similar argument, since $f(t) - f(t-1) > f(t_2^*) - f(f_2^* - 1) > f(t_2^* + 1) - f(f_2^*) > \frac{1+S}{S}[q_4(2) -$

$q_5(2)][E_g^l - E_r^l] > \frac{1+S}{S}[q_t(2) - q_{t+1}(2)][E_g^l - E_r^l] > (1-\beta)[q_t(1) - q_{t+1}(1)][E_g^m - E_r^m]$ for $3 \leq t < t_2^* + 1$, $w_{t+1}^m[q_{t+1}(1)] > w_t^m[q_t(1)]$; since $f(t+1) - f(t) < f(t_2^*+2) - f(t_2^*+1) < (1-\beta)[q_{\bar{t}}(2) - q_{\bar{t}+1}(2)][E_g^l - E_r^l] < \frac{1+S}{S}[q_{\bar{t}}(2) - q_{\bar{t}+1}(2)][E_g^l - E_r^l]$ for $t > t_2^* + 1$ and $f(t+1) - f(t) \to 0$ for $t \geq \bar{t}$, $w_{t+1}^m[q_{t+1}(1)] < w_t^m[q_t(1)]$. $\|$

APPENDIX C

APPENDIX FOR CHAPTER 3

C.1 Proof of Proposition 5 and 8

Consider a production unit with production efficiency V employing n young workers. In period 2, outside firms form expectations about old workers' ability denoted as θ^e. Due to the existence of the exogenous movers, the market expectation is free of the winner's curse. Since there are multiple firms, the standard result is that the market value of the worker is at her expected productivity, which equals to $V\theta^e$. Since the worker can produce more at her incumbent firm, her first-period employer has an incentive to match her market value.$\|$

C.2 Proof of Proposition 6

From (3.3) , since the RHS is unchanged, when V^E goes up, n goes up. $\|$

C.3 Proof of Proposition 7

From *Proposition* 6, since $V_1^E > V_2^E, V_1^L = V_2^L$, then $n_1 > n_2$. Since $\{\theta_i\}, i \in 1, \ldots, n$, follows uniform distribution, $f(\theta) = 1/(\theta_H - \theta_L), F(\theta) = (\theta - \theta_L)/(\theta_H - \theta_L)$. Prob.($\theta_i$ is the largest among n)=$F^n(\theta)$. Thus, $E(\theta_i|\theta_i$ is the largest among n$) = \int_{\theta_L}^{\theta_H} \theta \cdot nf(\theta)F^{n-1}(\theta)d\theta = (n\theta_H + \theta_L)/(n+1)$, which increases in n. Since $n_1 > n_2$, $E(\theta_i|\theta_i$ is the largest among $n_1) > E(\theta_i|\theta_i$ is the largest among $n_2)$. Thus, $W_1^E > W_2^E$. From the law of total expectation, $E(\theta) = E(\theta|\theta$ is not the largest among n)$\cdot P(\theta$ is not the largest among n)$+E(\theta|\theta$ is the largest among n)$\cdot P(\theta$ is the largest among n),$(\theta_H+\theta_L)/2 = E(\theta|\theta \text{ not largest})(1-1/n)+(n\theta_H+\theta_L)/(n+1)(1/n)$.$E(\theta|\theta \text{ not the largest}) = [n(\theta_H+\theta_L)+2\theta_L]/[2(n+1)]$, which also increases in n. Thus, the laborers' wage also increases, i.e., $W_1^L > W_2^L$. As for the wage difference, $\Delta W = (V^E - V^L)\theta^E(n) + V^L[\theta^E(n) - \theta^L(n)]$. We know $\theta^E(n)$ increases in n. $\theta^E(n) - \theta^L(n) = \frac{\theta_H - \theta_L}{2} \cdot \frac{n}{n+1}$, which also increases in n. Thus $\Delta W_1 > \Delta W_2$. $\|$

C.4 Proof of Proposition 9

Note that $\theta^E + (n-1)\theta^L = nE(\theta)$. Thus, the maximization problem at the two-layer firm can be re-write into

$$\underset{n}{Max} \quad nV^L E(\theta) + S(V^E - V^L)\theta^E(n) + SV_n^L nE(\theta) - n(\bar{U}_1 + \bar{U}_2).$$

I re-write as follows the three first-order conditions that pin down the equilibrium number of young workers in the three-layer firm and the two-layer firm.

$$S(V^E - V^L)\frac{\theta_H - \theta_L}{(N+1)^2} = (\bar{U}_1 + \bar{U}_2) - (1+S)V^L E(\theta) \quad \text{(C.1)}$$

$$S(\hat{V}^E - \hat{V}^M)\frac{\theta_H - \theta_L}{(\hat{m}+1)^2} = (\bar{U}_1 + \bar{U}_2) - (1+S)\hat{V}^M E(\theta) \quad \text{(C.2)}$$

$$\frac{1}{\hat{m}}S(\hat{V}^M - \hat{V}^L)\frac{\theta_H - \theta_L}{(\hat{n}+1)^2} = (\bar{U}_1 + \bar{U}_2) - (1+S)\hat{V}^L E(\theta) \quad \text{(C.3)}$$

From (C.3) over (C.2) , $\frac{\hat{V}^M - \hat{V}^L}{\hat{V}^E - \hat{V}^M} \cdot \frac{(\hat{m}+1)^2}{\hat{m}(\hat{n}+1)^2} = \frac{(\bar{U}_1+\bar{U}_2)-(1+S)\hat{V}^L E(\theta)}{(\bar{U}_1+\bar{U}_2)-(1+S)\hat{V}^M E(\theta)}$. Then the RHS>1. Suppose $\hat{n} > \hat{m}$. Since $\frac{\hat{V}^M - \hat{V}^L}{\hat{V}^E - \hat{V}^M} < 1, \frac{(\hat{m}+1)^2}{\hat{m}(\hat{n}+1)^2} < 1$, the $LHS < 1$. Contradiction. Thus, $\hat{n} < \hat{m}$.$\|$

C.5 Proof of Proposition 10

From (C.2),

$$S[(\hat{V}^E - \hat{V}^M) - (\hat{V}^M - \hat{V}^L)]\frac{\theta_H - \theta_L}{(\hat{m}+1)^2} = (\bar{U}_1 + \bar{U}_2) - (1+S)\hat{V}^M E(\theta) - S(\hat{V}^M - \hat{V}^L)\frac{\theta_H - \theta_L}{(\hat{m}+1)^2}. \quad (C.4)$$

From (C.1) over (C.5), we have

$$\frac{V^E - V^L}{\hat{V}^E - \hat{V}^L}\frac{(\hat{m}+1)^2}{(N+1)^2} = \frac{(\bar{U}_1 + \bar{U}_2) - (1+S)V^L E(\theta)}{(\bar{U}_1 + \bar{U}_2) - (1+S)\hat{V}^M E(\theta) - S(\hat{V}^M - \hat{V}^L)\frac{\theta_H - \theta_L}{(\hat{m}+1)^2}} > 1.$$

Since $\frac{(\hat{m}+1)^2}{(N+1)^2} < 1$, $\frac{V^E - V^L}{\hat{V}^E - \hat{V}^L} > 1, i.e., V^E > \hat{V}^E$. The CEO position at the flatter firm is more efficient. ||

C.6 Proof of Proposition 11

$W^E = V^L\theta^E(N), W^L = V^L\theta^L(N), \hat{W}_P^E = \hat{V}^L\hat{\theta}_P^E(\hat{m}), \hat{W}_N^L = \hat{V}^L[\frac{1}{\hat{m}}\hat{\theta}_N^L(\hat{n}) + (1-\frac{1}{\hat{m}})E(\theta)]$. From Proposition 10,$V^E > \hat{V}^E$. From Proposition 7,$\theta^E(N) > \hat{\theta}_P^E(\hat{m})$. Let's consider $A := \theta^L(N)$ and $B := \frac{1}{\hat{m}}\hat{\theta}_N^L(\hat{n}) + (1-\frac{1}{\hat{m}})E(\theta)$. We know $N = \hat{N} = \hat{m}(\hat{n}+1)$. Then $A := \frac{\hat{m}(\hat{n}+1)(\theta_H+\theta_L)+2\theta_L}{2[\hat{m}(\hat{n}+1)+1]}$. $B := \frac{\hat{m}(\hat{n}+1)(\theta_H+\theta_L)+2\theta_L-(\theta_H+\theta_L)}{2[\hat{m}(\hat{n}+1)+1]-2}$. Since $A < \frac{\theta_H+\theta_L}{2}$, then $A > B$. That is, the expected ability of a non-promoted laborer in a two-layer firm is higher than the expected ability of a non-promoted laborer in a three-layer firm. Therefore, if $\hat{N} = N, \hat{V}^L = V^L$, then $W^E > \hat{W}_P^E, W^L > \hat{W}_N^L$.

From the proof of Proposition 7, the difference in the two expectations $E(\theta|\theta \text{ is the largest in } n) - E(\theta|\theta \text{ is not the largest in } n) = \frac{\theta_H-\theta_L}{2}\cdot\frac{n}{n+1}$ increases in n. $\Delta W = V^L[\theta^E(N) - \theta^L(N)] = V^L\cdot\frac{\theta_H-\theta_L}{2}\cdot\frac{N}{N+1}$. $\Delta\hat{W} = \hat{V}^L[\hat{\theta}_P^E(\hat{m}) - \frac{1}{\hat{m}}\hat{\theta}_N^L(\hat{n}) - \frac{\hat{m}-1}{\hat{m}}\frac{\theta_H+\theta_L}{2}]$. Let $C := \hat{\theta}_P^E(\hat{m}) - \frac{1}{\hat{m}}\hat{\theta}_N^L(\hat{n}) - \frac{\hat{m}-1}{\hat{m}}\frac{\theta_H+\theta_L}{2} < \frac{\hat{m}^2-\hat{m}+1}{\hat{m}(\hat{m}+1)}\cdot\frac{\theta_H-\theta_L}{2} = [1 - \frac{2\hat{m}-1}{\hat{m}(\hat{m}+1)}]\frac{\theta_H-\theta_L}{2}$. Let

$D := \frac{N}{N+1} \cdot \frac{\theta_H - \theta_L}{2} = \frac{\hat{m}(\hat{n}+1)}{\hat{m}(\hat{n}+1)+1} \cdot \frac{\theta_H - \theta_L}{2} = [1 - \frac{1}{\hat{m}(\hat{n}+1)+1}] \frac{\theta_H - \theta_L}{2}$. We can see that $D > C$ as long as $\hat{m}, \hat{n} > 1$. Therefore, $\Delta W > \Delta \hat{W}$. $\|$

BIBLIOGRAPHY

[1] Abowd, J. M. (1990). Does performance-based managerial compensation affect corporate performance? Industrial and Labor Relations Review, 52S-73S.

[2] Acemoglu, D., and Pischke, J. S. (1998). Why do firms train? theory and evidence. Quarterly Journal of Economics, 113(1), 79-118.

[3] Acemoglu, D., and Pischke, J. S. (1999). The structure of wages and investment in general training. The Journal of Political Economy, 107(3), 539-572.

[4] Akerlof, G. (1970). The market for lemons. Quarterly Journal of Economics, 84(3), 488-500.

[5] Alchian, A. A., and Demsetz, H. (1972). Production, information costs, and economic organization. The American Economic Review, 62(5), 777-795.

[6] Altonji, J. G., and Blank, R. M. (1999). Race and gender in the labor market. Handbook of labor economics (Orley Ashenfelter and David Card ed., pp. 3143-3259) Elsevier.

[7] Altonji, J. G., and Paxson, C. H. (1988). Labor supply preferences, hours constraints, and hours-wage trade-offs. Journal of Labor Economics, 6(2), 254-276.

[8] Altonji, J. G., and Pierret, C. R. (2001). Employer learning and statistical discrimination. The Quarterly Journal of Economics, 116(1), 313-350.

[9] Altonji, J. G., and Shakotko, R. A. (1987). Do wages rise with job seniority? The Review of Economic Studies, 54(3), 437-459.

[10] Altonji, J. G., and Williams, N. (2005). Do wages rise with job seniority? A reassessment. Industrial and Labor Relations Review, 58(3), 370-397.

[11] Ariga, K. (2006). Horizontal transfer, vertical promotion, and evolution of firm organization. Journal of the Japanese and International Economies, 20(1), 20-49.

[12] Autor, D., and Handel, M. J. (2013). Putting tasks to the test: Human capital, job tasks, and wages. Journal of Labor Economics, 31(2 Part 2), S59-S96.

[13] Baker, G. P. (1992). Incentive contracts and performance measurement. The Journal of Political Economy, 100(3), 598-614.

[14] Baker, G., and Holmstrom, B. (1995). Internal labor markets: Too many theories, too few facts. The American Economic Review, 85(2), 255-259.

[15] Baker, G., Gibbons, R., and Murphy, K. J. (2002). Relational contracts and the theory of the firm. Quarterly Journal of Economics, 117(1), 39-84.

[16] Baker, G., Gibbs, M., and Holmstrom, B. (1994a). The wage policy of a firm. The Quarterly Journal of Economics, 109(4), 921-955.

[17] Baker, G., Gibbs, M., and Holmstrom, B. (1994b). The internal economics of the firm: Evidence from personnel data. The Quarterly Journal of Economics, 109(4), 881.

[18] Balmaceda, Felipe, Task-Specific Training and Job Design (2006). working paper.

[19] Barron, J. M., Berger, M. C., and Black, D. A. (2006). Selective counteroffers. Journal of Labor Economics, 24(3), 385409.

[20] Bartel, A. P., and Sicherman, N. (1998). Technological Change and the Skill Acquisition of Young Workers. Journal of Labor Economics, 16(4), 718-55

[21] Bauer, T. K., and Bender, S. (2001). Flexible Work Systems and the Structure of Wages: Evidence from Matched Employer-Employee Data, IZA working paper.

[22] Beaudry, P., and DiNardo, J. (1991). The effect of implicit contracts on the movement of wages over the business cycle: Evidence from micro data. The Journal of Political Economy, 99(4), 665-688.

[23] Becker, G. (1962). Investment in human capital: A theoretical analysis. The Journal of Political Economy,70(5), 9-49.

[24] Becker, G. (1964). Human capital: A theoretical and empirical analysis, with special reference to education. New York: National Bureau of Economic Research.

[25] Belzil, C., and Bognanno, M. (2008). Promotions, demotions, halo effects, and the earnings dynamics of american executives. Journal of Labor Economics, 26(2), 287-310.

[26] Belzil, C., Bognanno, M., and Poinas, F. (2012). Promotion determinants in corporate hierarchies: An examination of fast tracks and functional area.

[27] Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. The Journal of Political Economy, 75(4), 352-365.

[28] Berk, J. B., Stanton, R. and Zechner, J. (2010), Human Capital, Bankruptcy, and Capital Structure. The Journal of Finance, 65(3), 891-926.

[29] Bernhardt, D. (1995). Strategic promotion and compensation. The Review of Economic Studies, 62(2), 315-339.

[30] Bernhardt, D., and Scoones, D. (1993). Promotion, turnover, and preemptive wage offers. The American Economic Review, 83(4), 771-791.

[31] Bloom, N., Sadun, R., and Van Reenen, J. (2010). Does product market competition lead firms to decentralize? The American Economic Review, , 434-438.

[32] Bognanno, M. L. (2001). Corporate tournaments. Journal of Labor Economics, 19(2), 290-315.

[33] Bognanno, M., and Melero, E. (2012). Promotion signals, age and education.

[34] Bresnahan, T. F., Brynjolfsson, E., and Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. The Quarterly Journal of Economics, 117(1), 339-376.

[35] Brown, J. N. (1989). Why do wages increase with tenure? on-the-job training and life-cycle wage growth observed within firms. The American Economic Review, 79(5), 971-991.

[36] Bull, C. (1987). The existence of self-enforcing implicit contracts. The Quarterly Journal of Economics, 102(1), 147-159.

[37] Caliendo, L., Monte, F., and Rossi-Hansberg, E. (2012). The anatomy of French production hierarchies (No. w18259). National Bureau of Economic Research.

[38] Calvo, G. A., and Wellisz, S. (1979). Hierarchy, ability, and income distribution. The Journal of Political Economy, 87(5), 991.

[39] Cappelli, P., and Hamori, M. (2004). The path to the top: Changes in the attributes and careers of corporate executives, 1980-2001 (No. w10507). National Bureau of Economic Research.

[40] Card, D., and Hyslop, D. (1997). Does inflation "grease the wheels of the labor market"? In C. Romer, and D. Rober (Eds.), Reducing inflation: Motivation and strategy (pp. 71-114). Chicago,IL: University of Chicago Press.

[41] Carmichael, L. (1983). Firm-specific human capital and promotion ladders. The Bell Journal of Economics, 14(1) , 251-258.

[42] Cassidy, H., DeVaro, J., and Kauhanen, A. (2012). The Signaling Role of Promotions: New Evidence from European Firms.working paper.

[43] Cellini, S. R., Ferreira, F., and Rothstein, J. (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design. The Quarterly Journal of Economics, 125(1), 215-261.

[44] Chang, C., and Wang, Y. (1996). Human capital investment under asymmetric information: The pigovian conjecture revisited. Journal of Labor Economics, 14(3), 505-519.

[45] Chiappori, P. A., Salanie, B., and Valentin, J. (1999). Early starters versus late beginners. The Journal of Political Economy, 107(4), 731-760.

[46] Clemens, A. (2012). Position-specific promotion rates and the Fast track effect. Research in Labor Economics, 36, 77-107.

[47] Colombo, M. G., and Delmastro, M. (1999). Some stylized facts on organization and its evolution. Journal of Economic Behavior and Organization, 40(3), 255-274.

[48] Colombo, M. G., and Delmastro, M. (2008). The economics of organizational design: Theoretical insights and empirical evidence. Houndmills, UK: Palgrave MacMillan.

[49] Demougin, D., and Siow, A. (1994). Careers in ongoing hierarchies. The American Economic Review, 84(5) , 1261-1277.

[50] DeVaro, J., and Waldman, M. (2012). The signaling role of promotions: Further theory and empirical evidence. Journal of Labor Economics, 30(1), 91-147.

[51] Devereux, P. J., and Hart, R. A. (2006). Real wage cyclicality of job non-movers, within-company job movers, and between-company job movers. Industrial and Labor Relations Review, , 105-119.

[52] Devereux, P. J., Hart, R. A., and Roberts, E. (2013). Job spells, employer spells, and wage returns to tenure.

[53] Dohmen, T. J., Kriechel, B., and Pfann, G. A. (2004). Monkey bars and ladders: The importance of lateral and vertical job mobility in internal labor market careers. Journal of Population Economics, 17(2), 193-228.

[54] Eckstein, Z., and Wolpin, K. I. (1989). Dynamic labour force participation of married women and endogenous work experience. The Review of Economic Studies, 56(3), 375-390.

[55] Ekinci, E. Bonus Payments in a Corporate Hierarchy: Theory and Evidence. Cornell University, Mimeo, 2011.

[56] Eriksson, T., and Ortega, J. (2006). The adoption of job rotaton: Testing the theories. Industrial and Labor Relations Review, 59(4), 653-666.

[57] Eriksson, T., and Werwatz, A. (2005). The prevalence of internal labour market-sNew evidence from panel data. International Journal of Economic Research, 2(2), 105-124.

[58] Fairburn, J. A., and Malcomson, J. M. (2001). Performance, promotion, and the peter principle. Review of Economic Studies, 68(1), 45-66.

[59] Farber, H. S., and Gibbons, R. (1996). Learning and wage dynamics. The Quarterly Journal of Economics, 111(4), 1007.

[60] Frederiksen, A., and Kato, T.(2011) Human Capital and Career Success: Evidence from Linked Employer-Employee Data. IZA working papers (No.5764).

[61] Friebel, G., and Raith, M. (2013). Managers, Training, and Internal Labor Markets. Available at SSRN 2363492.

[62] Gabarro, J. J. (1987). The dynamics of taking charge Harvard Business Press.

[63] Garicano, L. (2000). Hierarchies and the Organization of Knowledge in Production. Journal of Political Economy, 108(5), 874-904.

[64] Garicano, L. (2000). Hierarchies and the organization of knowledge in production. The Journal of Political Economy, 108(5), 874-904.

[65] Garicano, L., and Rossi-Hansberg, E. (2006). Organization and inequality in a knowledge economy. Quarterly Journal of Economics, 121(4), 1383-1435.

[66] Gathmann, C., and Schonberg, U. (2010). How general is human capital? A task-based approach. Journal of Labor Economics, 28(1)

[67] Gergaud, O., Smeets, V., and Warzynski, F. (2012). Learning by Cooking and Reputation Building: A French Recipe to Become a Top Chef. Working paper.

[68] Ghosh, S., and Waldman, M. (2010). Standard promotion practices versus up-or-out contracts. The Rand Journal of Economics, 41(2), 301-325.

[69] Gibbons, R. (1997). Incentives and careers in organizations. In D. Kreps, and K. Wallis (Eds.), Advances in economics and econometrics: Theory and applications: Seventh world congress (eds. ed., pp. 1-37). New York: Cambridge University Press.

[70] Gibbons, R. (2005). Four formal (izable) theories of the firm? Journal of Economic Behavior and Organization, 58(2), 200-245.

[71] Gibbons, R., and Katz, L. F. (1991). Layoffs and lemons. Journal of Labor Economics, 9(4) , 351-380.

[72] Gibbons, R., and Murphy, K. J. (1992). Optimal incentive contracts in the presence of career concerns: Theory and evidence. The Journal of Political Economy, 100(3), 468-505.

[73] Gibbons, R., and Waldman, M. (1999a). A theory of wage and promotion dynamics inside firms. Quarterly Journal of Economics, 114(4), 1321-1358.

[74] Gibbons, R., and Waldman, M. (1999a). Careers in organizations: Theory and evidence. In O. C. Ashenfelter, and D. Card (Eds.), Handbook of labor economics, vol.3 (Orley C. Ashenfelter and David Card ed., pp. 2373-2437). Amsterdam: North-Holland:

[75] Gibbons, R., and Waldman, M. (2004). Task-specific human capital. The American Economic Review, 94(2), 203-207.

[76] Gibbons, R., and Waldman, M. (2006). Enriching a theory of wage and promotion dynamics inside firms. Journal of Labor Economics, 24(1), 59-107.

[77] Gibbons, R., Katz, L. F., Lemieux, T., and Parent, D. (2005). Comparative advantage, learning, and sectoral wage determination. Journal of Labor Economics, 23(4), 681-724.

[78] Gibbs, M. (1995). Incentive compensation in a corporate hierarchy. Journal of Accounting and Economics, 19(2-3), 247-277.

[79] Gibbs, M., and Hendricks, W. (2004). Do formal salary systems really matter. Industrial and Labor Relations Review, 58(1), 71-93.

[80] Golan, L. (2005). Counteroffers and Efficiency in Labor Markets with Asymmetric Information, Journal of Labor Economics, 23(2), 373-393.

[81] Grant, M. R. (2005). General electric: Life after jack. In M. R. Grant (Ed.), Cases to accompany contemporary strategy analysis (Fifth ed., pp. 336-353). United Kindom: Blackwell Publishing.

[82] Greenwald, B. C. (1986). Adverse selection in the labour market. The Review of Economic Studies, 53(3), 325-347.

[83] Guadalupe, M., and Wulf, J. (2010). The flattening firm and product market competition: The effect of trade liberalization on corporate hierarchies. American Economic Journal: Applied Economics, 2(4), 105-127.

[84] Haltiwanger, J., and Waldman, M. (1986). Insurance and labor market contracting: An analysis of the capital market assumption. Journal of Labor Economics, 355-375.

[85] Harris, M., and Holmstrom, B. (1982). A theory of wage dynamics. The Review of Economic Studies, 49(3), 315-333.

[86] Hlmstrom, B. (1979). Moral hazard and observability. The Bell Journal of Economics, 10(1), 74-91.

[87] Holmstrom, B., and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. Journal of Law, Economics, and Organization , Vol. 7, Special Issue: [Papers from the Conference on the New Science of Organization, January 1991] (1991), 24-52

[88] Imai, S., and Keane, M. P. (2004). Intertemporal labor supply and human capital accumulation*. International Economic Review, 45(2), 601-641.

[89] Kahn, C., and Huberman, G. (1988). Two-sided uncertainty and up-or-out contracts. Journal of Labor Economics, 6(4), 423-444.

[90] Kambourov, G., and Manovskii, I. (2009). Occupational specificity of human capital. International Economic Review, 50(1), 63-115.

[91] Kandel, E., and Lazear, E. P. (1992). Peer pressure and partnerships. The Journal of Political Economy, 100(4), 801-817.

[92] Kauhanen, A., and Napari, S. (2012). Career and wage dynamics: Evidence from linked employer-employee data. Research in Labor Economics, 36, 35-76.

[93] Keane, M. P., and Wolpin, K. I. (1997). The career decisions of young men. Journal of Political Economy, 105(3), 473-522.

[94] Kwon, I. (2006). Incentives, wages, and promotions: Theory and evidence. The Rand Journal of Economics, 37(1), 100-120.

[95] Kwon, I., Milgrom, E. M., and Hwang, S. (2010). Cohort effects in promotions and wages evidence from sweden and the united states. Journal of Human Resources, 45(3), 772-808.

[96] Lange, F. (2007). The speed of employer learning. Journal of Labor Economics, 25(1), 1-35

[97] Lazear, E. P. (1979). Why is there mandatory retirement? The Journal of Political Economy, 87(6), 1261-1284.

[98] Lazear, E. P. (1986). Raids and Offer-Matching, Research in Labor Economics,ed. Ronald Ehrenberg, Vol. 8, JAI Press, Greenwich, CT. 141-165

[99] Lazear, E. P. (1989). Pay equality and industrial politics. The Journal of Political Economy, 97(3), 561-580.

[100] Lazear, E. P. (2009). Firm-specific human capital: A skill-weights approach. Journal of Political Economy, 117(5), 914-940.

[101] Lazear, E. P., and Rosen, S. (1981). Rank order tournaments as optimum labor contracts. Journal of Political Economy, 89(5), 841-864.

[102] Li, F., and Tian, C. (2013). Directed search and job rotation. Journal of Economic Theory.148,1268-1281.

[103] Ljungqvist, L., and Sargent, T. J. (2004). Recursive macroeconomic theory. The MIT press.

[104] Lluis, S. (2005). The role of comparative advantage and learning in wage dynamics and intrafirm mobility: Evidence from germany. Journal of Labor Economics, 23(4), 725-767.

[105] MacDonald, G. M. (1982). A market equilibrium theory of job assignment and sequential accumulation of information. The American Economic Review, 72(5), 1038-1055.

[106] MacDonald, G. M. (1988a). The economics of rising stars. The American Economic Review, 78(1), 155-166.

[107] MacDonald, G. M. (1988b). Job mobility in market equilibrium. The Review of Economic Studies, 55(1), 153-168.

[108] MacLeod, W. B. (2007). Reputations, relationships, and contract enforcement. Journal of Economic Literature, XLV(September), 595-628.

[109] MacLeod, W. B., and Malcomson, J. M. (1988). Reputation and hierarchy in dynamic models of employment. The Journal of Political Economy, 96(4), 832-854.

[110] McCall, M., Lombardo, M., and Morrison, Ann. (1988). Lessons of experience: How successful executives develop on the job. Simon and Schuster.

[111] McCue, K. (1996). Promotions and wage growth. Journal of Labor Economics, 14(2) , 175-209.

[112] McLaughlin, K. J. (1994). Rigid wages? Journal of Monetary Economics, 34(3), 383-414.

[113] Medoff, J. L., and Abraham, K. G. (1980). Experience, performance, and earnings. The Quarterly Journal of Economics, 95(4), 703-736.

[114] Meyer, M. A. (1992). Biased contests and moral hazard: Implications for career profiles. Annales D'Economie Et De Statistique, , 165-187.

[115] Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. The Bell Journal of Economics, 12(2), 380-391.

[116] Milgrom, P. R. (1988). Employment contracts, influence activities, and efficient organization design. The Journal of Political Economy, 96(1), 42-60.

[117] Milgrom, P., and Oster, S. (1987). Job discrimination, market forces, and the invisibility hypothesis. The Quarterly Journal of Economics, 102(3), 453-476.

[118] Mincer, J. (1958). Investment in human capital and personal income distribution. The Journal of Political Economy, 66(4), 281-302.

[119] Mincer, J. (1962). On-the-job training: Costs, returns, and some implications. The Journal of Political Economy, 70(5), 50-79.

[120] Mintzberg, H. (1973). The nature of managerial work. Englewood Cliffs, NJ: Pretice-Hall. Morgan W. Mccall, Michael

[121] Murphy, K. J., and Zabojnik, J. (2004). CEO pay and appointments: A market-based explanation for recent trends. The American Economic Review, 94(2), 192-196.

[122] Murphy, K., and Zabojnik, J. (2007). Managerial capital and the market for CEOs. Available at SSRN 984376

[123] Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. Journal of Labor Economics, 13(4), 653-677.

[124] Neal, D. (1999). The complexity of job mobility among young men. Journal of Labor Economics, 17(2), 237-261

[125] O'Keeffe, M., Viscusi, W. K., and Zeckhauser, R. J. (1984). Economic contests: Comparative reward schemes. Journal of Labor Economics, 2(1), 27-56.

[126] Ortega, J. (2001). Job rotation as a learning mechanism. Management Science, 47(10), 1361-1370.

[127] Owan, H. (2004). Promotion, turnover, earnings, and FirmSponsored training. Journal of Labor Economics, 22(4), 955-978.

[128] Parent, D. (2000). Industry-specific capital and the wage profile: Evidence from the national longitudinal survey of youth and the panel study of income dynamics. Journal of Labor Economics, 18(2), 306-323.

[129] Pavan, R. (2011). Career choice and wage growth. Journal of Labor Economics, 29(3), 549-587.

[130] Pinkston, J. C. (2009). A model of asymmetric employer learning with testable implications. Review of Economic Studies, 76(1), 367-394.

[131] Podolny, J. M., and Baron, J. N. (1997). Resources and relationships: Social networks and mobility in the workplace. American Sociological Review, , 673-693.

[132] Postel-Vinay, F., and Robin, J. M. (2004). To match or not to match?:: Optimal wage policy with endogenous worker search intensity. Review of Economic Dynamics, 7(2), 297-330.

[133] Prendergast, C. (1993). The role of promotion in inducing specific human capital acquisition. The Quarterly Journal of Economics, 108(2), 523-534.

[134] Prendergast, C., and Topel, R. H. (1996). Faboritism in organizations. Journal of Political Economy, 104(5), 958-978.

[135] Qian, Y. (1994). Incentives and loss of control in an optimal hierarchy. The Review of Economic Studies, 61(3), 527-544.

[136] Rajan, R. G., and Wulf, J. (2006). The flattening firm: Evidence from panel data on the changing nature of corporate hierarchies. The Review of Economics and Statistics, 88(4), 759-773.

[137] Rodrguez, G., and Elo, I. (2003). Intra-class correlation in random-effects models for binary data. The Stata Journal, 3(1), 32-46.

[138] Rogerson, W. P. (1985). The first-order approach to principal-agent problems. Econometrica, 53(6), 1357-1367.

[139] Rosen, S. (1982). Authority, control, and the distribution of earnings. The Bell Journal of Economics, 13(2), 311-323.

[140] Rosen, S. (1986). Prizes and incentives in elimination tournaments. The American Economic Review, 76(4), 701-715.

[141] Saari, L. M., Johnson, T. R., McLaughlin, S. D., and Zimmerle, D. M. (1988). A survey of management training and education practices in US companies. Personnel Psychology, 41(4), 731-743.

[142] Sanders, C., and Taber, C. (2012). Life-cycle wage growth and heterogeneous human capital. Annual Review of Economics [First Published Online as a Review in Advance on April 5, 2012], 4, 399-425.

[143] Sasaki, M., Takii, K., and Wan, J. (2012). Horizontal transfer and promotion: New evidence and an interpretation from the perspective of task-specific human capital. Working Papers,

[144] Sattinger, M. (1975). Comparative advantage and the distributions of earnings and abilities. Econometrica, 43(3), 455-468.

[145] Schonberg, U. (2007). Testing for asymmetric employer learning. Journal of Labor Economics, 25(4), 651-692.

[146] Schultz, T. W. (1961). Investment in human capital. The American Economic Review, 51(1), 1-17.

[147] Schulz, E., Chowdhury, S., and Van de Voort, D. (2013). Firm productivity moderated link between human capital and compensation: The significance of task-specific human capital. Human Resource Management, 52(3), 423-439.

[148] Seltzer, A., and Merrett, D. (2000). Personnel policies at the union bank of australia: Evidence from the 1888-1900 entry cohorts. Journal of Labor Economics, 18(4), 573-613.

[149] Smeets, V., Waldman, M., and Warzynski, F. (2013). Performance, career dynamics, and span of control. Working Papers,

[150] Spence, M. (1973). Job market signaling. The Quarterly Journal of Economics, 87(3), 355.

[151] Stokey, N. L., Lucas, R. E., and Prescott, E. C. (1989). Recursive methods in economic dynamics. Harvard University Press.

[152] Sullivan, P. (2010). Empirical evidence on occupation and industry specific human capital. Labour Economics, 17(3), 567-580.

[153] Topel, R. H., and Ward, M. P. (1992). Job mobility and the careers of young men. The Quarterly Journal of Economics, 107(2), 439-479.

[154] Treble, J., Van Gameren, E., Bridges, S., and Barmby, T. (2001). The internal

economics of the firm: Further evidence from personnel data. Labour Economics, 8(5), 531-552.

[155] Waldman, M. (1984a). Job assignments, signaling, and efficiency. The Rand Journal of Economics, 15(2), 255-267.

[156] Waldman, M. (1984b). Worker allocation, hierarchies and the wage distribution. The Review of Economic Studies, 51(1), 95-109.

[157] Waldman, M. (1990). Up-or-out contracts: A signaling perspective. Journal of Labor Economics, 8(2), 230-250.

[158] Waldman, M. (2003). Ex ante versus ex post optimal promotion rules: The case of internal promotion. Economic Inquiry, 41(1), 27-41.

[159] Waldman, M. (2012a). Theory and evidence in internal labor markets. In R. Gibbons, and J. Roberts (Eds.), Handbook of organizational economics. Princeton University Press.

[160] Waldman, M. (2012b). Classic promotion tournaments versus market-based tournaments. International Journal of Industrial Organization, (0), Available online 21 March 2012.

[161] Waldman, M. and Zax, O. (2014). An Exploration of the Promotion Signaling Distortion. Working Paper.

[162] Williamson, O. E. (1967). Hierarchical control and optimum firm size. The Journal of Political Economy, 75(2), 123-138.

[163] Zabojnik, J. (2012). Promotion tournaments in market equilibrium. Economic Theory, 51(1), 213-240.

[164] Zabojnik, J., and Bernhardt, D. (2001). Corporate tournaments, human capital acquisition, and the firm size-wage relation. The Review of Economic Studies, 68(3), 693-716.