

ESSAYS ON EMPLOYEE TURNOVER

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

Jonathan R Peterson

January 2011

© 2010 Jonathan R Peterson
ALL RIGHTS RESERVED

ESSAYS ON EMPLOYEE TURNOVER

Jonathan R Peterson, Ph.D.

Cornell University 2011

This dissertation is a collection of theoretical works discussing the relationship between various human resource policies and employee retention. I build my models on a turnover mechanism motivated by workers' private information about their feeling toward their current employer. Workers use this private information when deciding whether to continue working for that institution. This turnover mechanism provides interesting insights into various employment and compensation practices. This dissertation discusses the interactions of turnover with pay systems, promotion and firm sponsored training.

The first chapter I discuss the role of formal pay systems within firms. The role of rigid bureaucratic wage rules within organizations has been a puzzle for some time. This paper provides an explanation for the use of these formal pay systems using two intuitive yet somewhat novel labor market assumptions. First, turnover is induced through wages and private worker non pecuniary taste shocks. Second, wage offers by current employers are a signal of worker ability. In this model, such asymmetry induces even higher turnover compared to full information over worker ability. This increased turnover reduces the expected welfare of the worker. I show how formal pay systems such as pay scales and budgets can mitigate this problem and reduce inefficient turnover. I also provide testable predictions in order to distinguish the wage signaling model from the symmetric information model.

In the second chapter, I build an infinite-horizon turnover model to address the relationship between turnover, promotion and wages. Firms are deeply concerned

with the costs of employee turnover. However, traditional labor economic theory is ill equipped to justify this concern. I explain how firms capture rent from its continual relationships with an employee. Consistent with empirical studies, I find that an employee's turnover rate will decrease once promoted. This paper also generates new empirical predictions as well as other well established wage and employee turnover dynamics.

In the third chapter, I examine the role of employee bonding contracts on turnover efficiency. Turnover is generated by realizations of a private taste shock observed by the worker after a period of work with the employee. These shocks allow the firm to exercise a type of monopsonistic power over their current employees creating inefficiently high turnover in equilibrium. I show that if done correctly, employee bonding contracts such as pension or minimum employment terms can reduce this inefficiency. The bond must be written so that the manager setting the wage does not directly benefit from the worker quitting. This separation of the wage setter and the bond holder is necessary to generate the required efficiency improvements. A competitive labor market ensures that these efficiency gains are ultimately realized by the worker. Workers prefer firms that use devices to reduce turnover inefficiencies. I compare the results of this model to empirical work on pensions and tuition reimbursement plans.

BIOGRAPHICAL SKETCH

Jonathan Peterson was raised in the small town of Ephraim Utah by parents who had an acute interest in education. His father worked as an instructor and in administration at the local junior college and his mother was an early childhood educator. As a child and adolescent Jonathan's interests were varied, but he excelled in topics relating to the math and sciences as well as music. He studied jazz percussion at Snow College for a while taking time to serve a full time mission for the Church of Jesus Christ of Latter Day Saints.

It was also at Snow College that he met and married his wife Shari. While at Snow College he also discovered his love for the subject Economics which offered an appealing combination of mathematics and social science. He soon decided to focus his studies away from music and toward the economics and mathematics. He continued his education at The University of Utah where enjoyed wonderful instruction and support and earned dual BS degrees in Economics and Mathematics. Encouraged by his instructors he applied to the best economics PhD programs. He was delighted when accepted to Cornell where he was to develop his skills as an economist.

At Cornell, Jonathan worked hard to learn the traits of a good economist and to be a good husband and Father. He felt blessed by many wonderful mentors who freely gave precious advice and time. He was fortunate to work as a student of Michael Waldman who taught him to discipline his work to the more important aspects of a problem. His chosen fields of study within economics were Industrial Organization and Organizational Economics. After years of work, Jonathan earned his PhD in economics in August of 2010 anxious to continue working many new and interesting problems.

To my wife Shari,
with whom I share this journey.

ACKNOWLEDGEMENTS

I am grateful to Justin Johnson, Talia Bar and Heski Bar-Isaac as well as participants in seminars at Cornell and BYU for helpful comments and encouragement. I would like to especially thank Michael Waldman for his time, insights, and unfailing attention to detail. All remaining errors are my own.

TABLE OF CONTENTS

Biographical Sketch	iii
Dedication	iv
Acknowledgements	v
Table of Contents	vi
1 Asymmetric Information and Wage Signals: The Need for Formal Pay Systems	1
1.1 Introduction	1
1.2 Literature Review	5
1.3 The Model	9
1.3.1 Symmetric Information about ability	9
1.3.2 Wages as Signals	15
1.4 Pay Systems	22
1.4.1 Single Wages	24
1.4.2 Pay Scales	27
1.4.3 Budgets	32
1.5 Discussion	39
1.6 Conclusion	40
2 Turnover and Promotion	42
2.1 Introduction	42
2.2 Literature review	46
2.3 The Turnover Mechanism	50
2.3.1 Simple Model	50
2.3.2 Wage gains at voluntary turnover	55
2.3.3 Cost of Turnover	55
2.3.4 Efficient turnover	56
2.3.5 Turnover rate dynamics	57
2.3.6 Switching costs	58
2.4 Job assignment and training costs	59
2.4.1 Model	60
2.4.2 Equilibrium	62
2.4.3 Turnover rate dynamics	68
2.4.4 Wage Dynamics	68
2.4.5 Switching Cost and Promotion	70
2.5 Extension	71
2.6 Discussion	73
2.6.1 Promotion and Turnover	73
2.6.2 Wage premium at promotion	74
2.7 Conclusion	76

3	Employee Bonding and Turnover Efficiency	77
3.1	Introduction	77
3.2	Literature Review	79
3.3	The Model	85
3.3.1	No Employee Bonding	85
3.4	Bonding	91
3.4.1	Managers holding the bond.	92
3.4.2	Firms holding the bond	93
3.5	General Skills Training	97
3.5.1	Worker Paid Training	98
3.5.2	Firm Sponsored Training	99
3.6	Firm Specific Training and Pensions	102
3.7	Conclusion	105
A	Asymmetric Information and Wage Signals:	107
B	Turnover and Promotion	114
C	Employee Bonding and Turnover Efficiency	127
	Bibliography	130

CHAPTER 1

ASYMMETRIC INFORMATION AND WAGE SIGNALS: THE NEED FOR FORMAL PAY SYSTEMS

1.1 Introduction

The existence and prevalence of formal pay systems has been a puzzle among labor economists for quite some time (see Doeringer and Piore 1971). In their seminal paper quoted above, Baker, Gibbs and Holmstrom (1994) discuss and puzzle over the use of pay scales in the firm that they study. While there have been many studies that have documented the various particulars of such systems, there are few economic theories that provide a reason for their use. I argue that such restrictions have an important role in improving efficiency through a reduction in employee turnover. To my knowledge this is the first paper to consider turnover reduction as a motivation for such rules.

I develop a two period labor turnover model to examine the usefulness of pay restrictions. I assume perfect competition between firms for workers. Any expected future profits are given to the worker at the beginning of the relationship. Therefore any policy that improves efficiency improves worker utility and is preferred by the workers..

The need for pay restrictions arises under specific information assumptions. These assumptions are asymmetric information about worker ability, wages as signals, and private worker taste shocks. In this case, the equilibrium wage profile causes higher turnover rates and lower profits than under symmetric information. I show that firms are able to reduce or eliminate these inefficiencies through formal

pay systems such as pay scales or budget restrictions.

In my model the threat of employee turnover plays a key role. The turnover mechanism I use is relatively unexplored, has appealing properties and provides important new insights on the relationship between the worker and firm. Similar to many other turnover models I assume that there is a match specific rent that is realized after a period. The difference between this model and standard turnover models is the information structure of the firm-worker match. In this model, the match specific taste shock is learned only by the employee after a period of work with his current firm.¹ These private taste shocks have a natural interpretation. They are the worker's attitudes toward the non pecuniary aspects of the job. Such non pecuniary aspects might include relationships with co workers, like or dislike of the geographic area or family concerns. This taste shock is considered by the worker when evaluating his quitting decision. This information asymmetry between worker and firm induces a trade-off between retention and ex post profit through the wage offer. Offering a lower wage to the worker increases the rent on the employee if he stays. But, it also reduces the chance that he will accept the offer.

The worker's information about his own preferences gives current employers monopsonistic power. Suppose the current employer offers a wage below marginal product. Even though the worker is getting an outside offer equal to marginal product the worker will accept the lower inside offer if his taste shock is high enough. Therefore this private shock provides the current firm positive profit in expectation. Since the current firm wages are below the outside wages the

¹Traditional turnover models also use match specific rents to drive turnover. However the firm always knows this match and is able to retain the worker when it is efficient. Here this match specific component is known only by the worker which generates a different wage setting decision.

resulting turnover rates are inefficiently high.

I analyze this model under two different informational assumptions regarding worker productivity. First, I assume that worker ability is common knowledge. I show that the resulting profit is an upper bound to any equilibrium where wages are a signal of ability. I then examine the model under the assumption that worker ability is only observed by the worker's current employer. I also assume that the wage offered by the worker's current firm is a signal of ability to the outside market.

I show that the equilibrium wage profile for the wage signaling model induces higher turnover probabilities than under symmetric information between firms. The excessive turnover is more extreme for workers of lower ability. The dual signaling and retention roles of wages are the source of this result. The current firm has an incentive to signal low ability through a lower wage offer. The outside market anticipates this and market expectations are less responsive to low wages. Therefore, the inside-outside wage gap is larger for lower ability workers.

Once I have shown the adverse effects of asymmetric information in this model I examine possible solutions. These solutions are various commitment devices such as formal pay systems. These systems allow the firm the ability to commit to future wage profiles before actually observing ability types. While these systems have differences their usefulness comes from their rigidity. By restricting the wage offers ex ante the firm reduces its capacity to signal worker ability in the future.

One particular system that I consider is pay scales. Here the firm commits to pay workers within a certain range of wages. This range of wages is set before observing worker ability. If the commitment is credible, this changes the signaled ability given the actual wage offer. Outside beliefs are directly tied to outside offers

and worker turnover. By restricting the possible wage offers, the firm can reduce its incentive to underpay and therefore decrease employee turnover in equilibrium.

The second pay system that I consider is the use of managers and budget restrictions. Here managers are given incentives to maximize expected output conditional on a budget set by the firm. While turnover is an issue for the managers there is no longer the same cost saving incentive beyond staying within budget. By evaluating the managers based on output rather than on profit the wage setters have no incentive to reduce costs below the budgeted amount. In this case there exists equilibria where the firm is able to achieve more efficient turnover rates.

I also analyze the effect of budget restrictions and pay scales under symmetric information. While such restrictions can help to mitigate the inefficiencies of wage signaling they can be harmful under symmetric information. Just as conventional wisdom suggests, strict budgetary restrictions on managers do not provide enough freedom for managers to set profit maximizing wage offers. Therefore I suggest that such pay systems are most likely to be observed in markets where worker ability is not common knowledge and compensation is observable.

The remainder of the paper is as follows. I review the relevant literature in Section 2. I set up and analyze the symmetric information and wage signaling models in Section 3. Section 4 examines the effect of restrictive pay systems under the different information assumptions. In Section 5 I discuss these results in response to classical adverse selection models. Section 6 is the conclusion.

1.2 Literature Review

From the early discussion in Doeringer and Piore (1971) economists have puzzled over the prevalence and rigidity of formal pay systems. These formal pay systems have been documented by a number of subsequent studies (ex. Baker et al 1994, Dohmen 2004). Gibbs and Hendricks (2004) describes these systems as such, "the firm used centralized policies to set salary levels and ranges, and determined how performance ratings were used to award raises and bonuses." While these systems may differ from firm to firm they have the common characteristic of being pre determined centrally and at least somewhat rigid. While there is a vast literature discussing job assignment, compensation and training of workers, the rigidity and purpose of this bureaucracy remains relatively unexplored.

So far, the work on formal pay rules within an organization has focused on improving performance by managers and employees. For example the theories on tournaments (ex. Lazear and Rosen 1981) and up-or-out contracts (ex. Kahn and Huberman 1981) rely on the firm's ability to commit to reward workers that exert high investments in human capital. Other theoretical work focuses on limiting the discretion of managers thus reducing employee influence activities (see Milgrom 1988, Milgrom and Roberts 1988a, 1988b) or favoritism (see Prendergast and Topel 1996).

Asymmetric information about worker ability has long been a central theme in personnel economics. One topic of particular interest is how asymmetric information about worker ability is advantageous to employers. Most of the theoretical work on this type of asymmetry revolves around firms establishing, maintaining, and capitalizing on this advantage (ex. Greenwald 1986).² There has been a great

²For example Acemoglu and Pichke (1998, 1999) show that adverse selection caused by infor-

deal of interest in the effect of signals on labor market outcomes. One of the most discussed signals in this literature is job assignment. Waldman (1984) was the first to show how this observable decision by the firm can affect the profitability of this information asymmetry. He shows why firms might distort the employee promotion decision to capture additional information rents. The idea of promotions as signals has been extended in many different ways (see Milgrom and Oster 1987, Ricart I Costa 1988, Waldman 1990, Bernhardt 1995, Golan 2005). Although the result that workers are promoted less than the efficient level is not fully robust (see Golan 2005), the result that information asymmetry is exploited by the firm is a common finding in this literature.³ This result however does not hold in the current paper due to the signaling role of the wage offer. Additional discussion of the difference between this literature and the current paper is provided in Section 1.5.

One important assumption in this paper is that wages are a signal of ability. The signaling aspect of wage offers has been largely ignored with the notable exception of Golan (2009).⁴ She uses a multi period bargaining model where productivity is known only by workers and current employers. She shows that high ability workers bargain with their employer to reduce their own wages early in their career in order to fully reveal their ability. Later in their career, symmetric information creates higher wages for more able workers. Higher ability workers experience faster wage growth even without any increase in productivity. Although wage signals play a key role in Golan's work as well as mine, the setup and results

mation asymmetries can induce firm sponsored employee training.

³ For example Milgrom and Oster (1987) show that workers with visible high ability are more likely to be promoted. Promoting workers with visible high ability reduces information rent by a smaller amount. DeVaro and Waldman (2006) provide evidence along these lines focusing on how promotion varies with the education level.

⁴In the asymmetric information models discussed above wages are not a signal of ability. Workers with common observables will be offered the same wages in equilibrium. Therefore the wage offer conveys no additional information.

of the paper are quite different. There is no turnover in her model. Also in her model there is no need for commitment devices such as pay systems.

In addition to the vast theoretical work there has been evidence of some general predictions of adverse selection models. Gibbons and Katz (1991) find that the wages of laid off workers are lower than those displaced after a plant closing where pre termination wages are identical (see also Kahn 2009).⁵ Doiron (1995) finds similar evidence using Canadian data. But Grund (1999) finds no such effect in his study of the German labor market. Stevens (1997) also finds that multiple displacements have a lasting effect on wages.

Uta Schonberg (2007) performs similar tests and finds evidence of asymmetric learning for college graduates and symmetric learning for workers with only a high school education. In addition to job assignment (ex. DeVaro and Waldman 2006) empirical economists have looked at previous job mobility (see Zhang 2007) or private signals (See Pinkston 2008) to test for these effects. Although not unanimous, the evidence is largely consistent with information asymmetry in a wide variety of markets.

In this paper, turnover plays a crucial role. There is a vast literature on various types of turnover mechanisms. The papers on job search and matching provide valuable insights into the turnover process (ex. Jovanovic 1979a, 1979b; Burdett 1978). In these classic models, turnover is a result of economic agents attempting to improve expected match quality. Imperfect information and learning gradually change worker and firm matches over time. While these models are able to generate

⁵This approach to testing for asymmetric information has critics who argue that this difference is driven by recall bias (see Song 2007) or differences in pre displacement wages (see Krashinsky 2002). Krashinsky argues that firms who lay off workers tend to be larger firms that pay higher wages in general. Therefore the large change in the wages by workers displaced by a layoff is driven by the high pre termination wage.

a number of empirical predictions they do not address the idea of employee turnover costs.⁶ Many such theories assume workers are paid their marginal product (ex. Jovanovic 1979a, 1979b). In this case there is no real loss to the firm when the worker leaves. In other turnover models workers are given a fixed proportion of their match specific rent (ex. Mortensen 1978). In this case firms do incur an economic loss when the worker switches employers. However a fixed sharing rule of the match specific rent may not be realistic. The ‘cost of employee turnover’ is not adequately addressed by these classical turnover models.⁷

The main difference between this paper and the classic turnover papers lies in the information assumption about match quality. In the current paper I assume that the quality of the match is known and experienced by the worker. In another chapter of this dissertation I use a similar assumption as I consider promotion timing, wages, and turnover (see chapter 2). In both papers the taste shock gives the current employer some monopsonistic power when choosing the wage offer. However, when workers decide to quit this induces an economic loss to the firm. I argue that this economic loss is the foundation for employee turnover costs. In chapter 2, I show that this dynamic monopsony power over current employees induces firms to create and maintain relationships with individual workers over time. I use this model to consider the promotion timing decision when promotion is costly to the firm and the firm experiences outside pressures. In chapter 3 I consider how employee bonding reduces turnover inefficiencies caused by managers setting

⁶For example, in their study of four different hotels Hinkin and Tracy (2000) estimate the turnover cost of a front-desk associate to be from 5 to 12 thousand dollars. Over half of this figure is non-explicit “productivity” costs. See also Cascio (1999) and Wasmuth and Davis (1983) for additional examples of costs of employee turnover.

⁷Under traditional theories, when counteroffers are possible, the firm will never allow a worker to leave before offering him a wage equal to his marginal product. If workers are paid their marginal product, turnover of a single worker will not affect the firm’s profitability. Under some restrictive assumptions, existing theory can generate endogenous turnover while firms pay a wage below marginal product. For example, a model with random outside wage offers that assumes away counteroffers may have this feature (see, Munasinghe and Flaherty 2005).

wages below marginal product. Aside from my work, a few papers use a similar turnover mechanism based on private worker utility (ex. Novos 1994, Acemoglu and Pischke 1998, Schonberg 2007, Ghosh 2007). Although none of them examine the signaling aspect of wages that is the core of this paper.

1.3 The Model

In this section I introduce the turnover mechanism that drives this model. I develop intuition on the source of turnover costs and the rents associated with attracting and retaining individual workers. In order to isolate the effects of the turnover mechanism I first assume that information on worker productivity is symmetric. This initial analysis provides a benchmark for later analysis.

1.3.1 Symmetric Information about ability

There is free entry into production. All firms are identical. The only input is labor and it is inelastically supplied each period. A worker's career lasts 2 periods. All firms and workers are perfectly patient. All workers have marginal products that are i.i.d. from a commonly known distribution F with support $[a_L, a_H]$. I will denote worker i 's marginal product each period as $a_i \in [a_L, a_H]$. In the first period all information about ability is unknown. After a period of work the worker's ability becomes common knowledge.

Worker i 's utility consists of wages, a taste shock and a switching cost $\sigma \geq 0$ (if incurred). The switching cost is incurred only when the worker switches employers. Worker i 's taste shock for firm k at time t is denoted as $\lambda_{i,t}^k$ and is private

information for the worker. All λ 's are drawn independently from a commonly known continuously differentiable distribution G . The distribution G has a zero mean and a continuously differentiable density g . The hazard rate $H(\cdot) = \frac{g(\cdot)}{1-G(\cdot)}$ is non-decreasing on its support. To ensure that turnover is sometimes efficient, I restrict the lower bound of the support of G to be $< -\sigma$. The worker only knows his second period taste shock for the firm he worked at in period 1. All other taste shocks are unknown. Since all firms are ex ante identical in the first period, the worker chooses the highest wage for his initial job. Perfect competition between potential employers ensures that second period expected profits are part of the worker's period 1 wage.

At the beginning of period 2 the current and outside firms offer wages to the worker. Given the current firm's wage offer the worker's utility if he stays with the firm in period 2 is,

$$u_2 = w_i + \lambda_i.$$

The worker incurs the same switching costs for all outside firms. If worker i enters the outside market, he accepts the highest outside firm wage offer \bar{w}_i . In this case his expected utility in period 2 is,

$$u_2 = \bar{w}_i - \sigma. \tag{1.1}$$

I assume that, if indifferent, the worker remains with the current firm. Therefore, the worker will stay with his current employer if $w_i + \lambda_i \geq \bar{w}_i - \sigma$. The worker will leave the current firm with probability,

$$\Pr(w_i + \lambda_i \geq \bar{w}_i - \sigma) = 1 - G(\bar{w}_i - w_i - \sigma). \tag{1.2}$$

The payoff to the firm is simply the marginal product of its employees minus the wage costs. A firm's wage decision for a given worker i , depends on the first

period employment relationship.⁸

The definition of equilibrium that I use is Perfect Bayesian Equilibrium. In the first period all firms are identical so the worker chooses the firm that offers the highest wage W . In the second period all firms learn the ability of the worker, but the value of λ is known only by the worker himself. All firms make wage offers and the worker chooses between offers. However the worker chooses his second period employer after observing his shock λ and all wage offers. The timing of the second period game is as follows.⁹

Stage 1: All firms observe ability and the current firm offers a wage.

Stage 2: Outside firms observe the current firm wage offer and offer wages.

Stage 3: Employee observes the taste shock for the current firm and chooses between offers.

Stage 4: Production, payment and utility experienced.

In equilibrium the outside market will always bid wages up until the outside firms are making zero expected profits.¹⁰ The current firm knows the worker's outside wage offer and chooses a wage to maximize profits. This gives us a

⁸We have not yet explicitly defined the outside option for the worker. It seems natural that wages should be greater than zero. In order to avoid uninteresting corner solutions or negative wages we impose the restriction that $a_l > x^*$ (where x^* uniquely solves $x^* \frac{g(x^* - \sigma)}{1 - G(x^* - \sigma)} = 1$). However, if we instead allow negative wages and assume that there is no outside option the results of the model are unchanged.

⁹In the benchmark model the timing is less important. Since workers enter the market with positive probability and ability is common knowledge, they will be offered their marginal product by the outside market regardless of the current firm's wage offer. However, to be fully consistent with the subsequent model we specify the timing as below for both models.

¹⁰We know this because the current firm will never offer a wage greater than marginal product. The lower bound of the distribution of the taste shock is less than minus the switching cost. Therefore there is a strictly positive probability that the worker will enter the outside market. Since the worker enters the market with a strictly positive probability, Bertrand like competition forces the wage offer to equal marginal product.

definition of equilibrium for the second period game.

Definition 1 *The equilibrium of this economy for worker i is a current firm wages offer w_i , outside wage offer \bar{w}_i such that,*

1. $w_i \in \arg \max_w (a_i - w) [1 - G(\bar{w}_i - w - \sigma)]$ (*Profit Maximization*)
2. $\bar{w}_i = a_i$ (*Zero Profit for Outside Firms*)

Since the worker has information about his taste for the current firm, the current firm has some monopsonistic power. It will offer a wage below marginal product and if the worker's taste shock is high enough, the worker will accept the offer. The expected profits for the current firm in the second period is,

$$V_i = \max_w (a_i - w) [1 - G(\bar{w}_i - w - \sigma)]. \quad (1.3)$$

This gives the firm positive expected profits in the second period. I now solve for the equilibrium of this game. Substituting the Zero Profit condition into the Profit Maximization condition I get,

$$V_i = \max_w (a_i - w) [1 - G(a_i - w - \sigma)] \quad (1.4)$$

$$= \max_{x_i} x_i [1 - G(x_i - \sigma)] \quad \text{where } x_{i,t} = a_i - w. \quad (1.5)$$

x is the ex post rent that a firm receives if the worker stays for the second period. However, increasing x increases turnover. Note that the optimal value of x is tied directly to the distribution of λ .

The first order condition of (1.5) in terms of x_i is,

$$x_i = \frac{1 - G(x_i - \sigma)}{g(x_i - \sigma)} = \frac{1}{H(x_i - \sigma)}. \quad (1.6)$$

Recall that the hazard rate $\frac{g}{1-G}$ is non-decreasing. The objective function is quasi-concave so the first order condition is necessary,¹¹ sufficient and has a unique solution. Let x^* solve the first order condition (1.6). Since λ 's are drawn i.i.d. for all ability types the equilibrium satisfies $x_i = x^*$ for all i . From (1.4) and (1.5) I know that V_i must also be the same for all i .

$$V_i = \max_w (a_i - w) [1 - G(\bar{w}_i - w - \sigma)] \quad (1.7)$$

$$= x^* [1 - G(x^* - \sigma)] = V^* \text{ for all } i. \quad (1.8)$$

This gives us the unique equilibrium for every realized ability.

Proposition 1 *The unique equilibrium of the above economy is as follows, where x^* and V^* are as defined above.*

$$w_i = a_i - x^* \quad (1.9)$$

$$\bar{w}_i = a_i \quad (1.10)$$

$$V_i = V^* \quad (1.11)$$

$$\text{Turnover probability} = G(x^* - \sigma) \quad \forall i. \quad (1.12)$$

Proof. above ■

Proposition 1 shows us that second period profit and turnover for any employee will be the same regardless of realized ability. The current firm will offer a wage that is exactly x^* less than marginal product. This uniform result comes from the i.i.d. assumption of the worker's taste shocks. If the worker stays with the current firm in period 2 the current firm earns a profit of x^* regardless of ability

¹¹ $x_{i,t} \leq 0$ cannot solve the maximization problem since positive profits are possible. Therefore the solution to the maximization problem must be on the open set $(0, \infty)$. If the maximum exists it must satisfy the first order condition.

type. Since the difference between current and outside wage offers is always x^* , the turnover probability is always $G(x^* - \sigma)$. The expected second period profit from employing worker i in the first period is $x^* [1 - G(x^* - \sigma)] = V^*$ regardless of the realization of ability.

I now discuss how the information asymmetry between the worker and firm causes turnover to be too high in equilibrium. If the current firm could observe the worker's taste shock, the firm would extract all rents from the shock if it is efficient for the worker to remain with the firm. If the shock is greater than the negative switching cost the inside firm would pay just enough to induce the worker to stay. If the shock is too low the firm will never offer a high enough wage and the worker will quit. However in this model the firm does not know the taste shock. It must offer a wage depending only on the worker's ability. It cannot extract all rents but gains rents in expectation by setting this wage below marginal product ($w_i = a_i - x^*$). The worker receives an outside wage offer of exactly marginal product. Therefore he will switch firms whenever $\lambda < x^* - \sigma$. The socially optimal turnover rule is to quit when $\lambda < -\sigma$. The current firm's exploitation of monopsonistic power results in inefficiently high turnover. I later show that this inefficiency is even greater when wage offers are a signal of ability.

From this simple model I have established a few insights into the relationship between the worker and the firm. The information on how the worker feels toward their current employer is valuable to that firm. Wages optimally set below marginal product means that firms experience an economic loss when workers quit.

1.3.2 Wages as Signals

So far I have examined the model under symmetric information. I now analyze the model when only the worker's current employer directly observes worker ability. I also assume the current firm's wage offer is observed by the outside market. It is possible now that the outside market infers something about the worker's ability from the current firm's wage offer. I show that the equilibrium wage profile induces higher turnover and lower second period profit than under symmetric information.

The assumptions of the model are the same as before except that now only the current firm has information about productivity. The strategies for the current firm are the possible wage offers depending on realized worker ability. The strategies for outside firms are the wage offers depending on beliefs. The beliefs of the outside market are conditional on the current firm wage offer. Let us denote the current firm's wage offer depending on ability as $w_i(a)$. I will denote the outside market's expectation of worker i 's ability as $b_i(w)$. For tractability I restrict beliefs to be the same for all firms.¹² Bertrand like competition results in the outside market always paying a wage equal to expected marginal product. Therefore $b_i(w)$ is also the outside market's wage offer if it observes the signal w from the current firm.

Given the belief function $b_i(w)$, the current firm's expected payoff when observing ability $a_i \in [a_L, a_H]$ and offering a wage $w_i(a_i)$ equals,

$$\max_w (a_i - w_i(a_i)) [1 - G(b_i(w_i(a_i)) - w_i(a_i) - \sigma)]. \quad (1.13)$$

¹²On the path of play this will naturally be satisfied in equilibrium. But off the path of play outside firms could have differing beliefs about worker ability. The game between outside firms would result in some expected maximal wage offer. Since the quitting decision is only determined by the inside vs. outside options, we assume away such technicalities.

I now define Perfect Bayesian Equilibrium for this model. A PBE is a wage policy $w_i(\cdot) : [a_H, a_L] \rightarrow \mathbb{R}$ and a set of beliefs $b_i(\cdot) : \mathbb{R} \rightarrow [a_H, a_L]$ that satisfy Condition 1 (Sequential Rationality) and Condition 2 (Consistency) as defined below.

Condition 1 (*Sequential Rationality*)

$$w_i(a_i) \in \arg \max_w (a_i - w) [1 - G(b_i(w) - w - \sigma)] \quad \forall a_i \in [a_l, a_h] \quad (1.14)$$

This condition guarantees that given the worker's realized ability and strategy of the outside market, the current firm is maximizing profits. Consistency of beliefs for this equilibrium can be a bit unclear since the distribution of ability types is continuous.¹³ However the following claim makes the resolution of this problem straightforward.

Claim 1 *Any wage profile that is part of a Perfect Bayesian Equilibrium must be weakly increasing with ability.*

Proof. In Appendix ■

This result that comes from the fact that firms with higher ability workers have more to lose from a worker quitting. If it is worth paying a higher wage to retain a low ability worker, it must also be worth paying the high wage to keep the high ability worker as well. Because wages are weakly increasing with ability, any wage offered in equilibrium is offered to a single ability type or a range of abilities. This result greatly simplifies the condition on the consistency of beliefs.

¹³In general it is unclear what the beliefs at a wage means when the set of ability types that offer that wage is not convex.

Condition 2 (*Consistency*) for any equilibrium current firm wage offer, $w_i(a_i)$, the outside offer, $b_i(w_i(a_i))$, must be the expected ability type of all workers offered $w_i(a_i)$, or the ability type of the single worker type offered $w_i(a_i)$.

Now that I have defined equilibrium in this setting I can find and characterize the equilibrium. The first type of equilibrium that I discuss is the perfectly separating equilibrium. This is when each ability type is offered a different wage.

It seems natural that higher ability workers are offered a higher wage. Although in equilibrium the relationship need only be weak. A wage profile that is strictly increasing with output might seem natural to avoid perverse incentives for workers to under perform.¹⁴ Even in this model with no worker effort, the fully separating equilibrium is interesting as it has similarities to the symmetric information outcome.¹⁵

Proposition 2 Let the function $\hat{a}(x)$ be defined as,

$$\hat{a}(x) = a_H + \int_{x^*}^x \left(1 - z \frac{g(z - \sigma)}{[1 - G(z - \sigma)]} \right) dz \quad \text{for } x > x^*. \quad (1.15)$$

The unique fully separating equilibrium is $w_i(a_i) = a_i - \hat{a}^{-1}(a_i)$ for all $a_i \in [a_L, a_H]$ with beliefs $b_i(w) = w_i^{-1}(w)$ for all $w \in [w_i(a_L), w_i(a_H)]$.¹⁶

Proposition 2 shows that there is an equilibrium to this game. Further, the fully separating equilibrium wage profile is unique. This equilibrium has the property

¹⁴That is, suppose output was a function of effort as well as ability. The same wage for multiple levels of output can induce some workers to exert inefficiently low effort.

¹⁵Most notably neither the fully separating equilibrium in the case of asymmetric information nor the symmetric information outcome have multiple ability types offered the same wage. Therefore by comparing the fully separating equilibrium to the full information outcome we can generate other predictions to distinguish these two models in the data.

¹⁶These equilibrium beliefs off of the path of play are not unique but one solution is $b_i(w) = a_H$ for all $w \notin [w_i(a_L), w_i(a_H)]$.

that turnover and expected second period profitability on the highest ability type will be the same as in the symmetric information case. Also, turnover for each ability type below a_H is higher than under symmetric information. Also second period profitability is lower for lower ability types.

I now discuss why the fully revealing wage and turnover rate for the high type must be the same as in the symmetric case. In the fully revealing case the equilibrium wage offered to the highest ability type will cause the outside market to offer a_H . Therefore under full revelation the upper bound on second period profit is the second period profits from the symmetric information case. Further since a_H is the maximum ability available, the outside market will never offer a wage above a_H . Therefore if the current firm offers the symmetric wage to the highest ability type the firm can attain at least the second period profit from the symmetric case. The wage offered to the highest ability type in equilibrium will be the same as in the symmetric case.

I now discuss the rest of the equilibrium. The current firm's wage offer directly affects the outside offers through the signaling role of the wage offer. Because workers act optimally, turnover is increasing in the difference between outside and inside wage offers ($b_i(w_i) - w_i$). This means that the turnover rate is a function of the current firm's wage offer. To show that turnover must be decreasing with the current firm's wage offer in equilibrium, I assume otherwise and find a contradiction. Expected profits are the product of ex post rents ($a_i - w$) and the probability that the worker stays ($1 - G(\cdot)$). A firm with a higher ability worker always has the option of paying the low ability equilibrium wage. If he does this, then it follows that ex post rents must be higher than if he paid the higher wage. If turnover is weakly lower for the lower wage, then profits associated with paying

the low wage are higher. This means that the inside firm would prefer to pay the low wage to the higher ability worker. This contradicts the result that more able workers are paid higher wages.

As noted above, the turnover rate of the highest ability type is the same as under symmetric information. Also, turnover rates on all other types must be higher than under symmetric information. Therefore, the inefficiencies associated with excess turnover are greater in this equilibrium than under symmetric information.

For a given employee lowering the wage offer has three effects on expected profit. The first is the obvious effect of lowering the costs to the firm. The second effect is that a lower wage offer signals to the outside market that the worker is of lower ability. This negative signal reduces his outside option, decreases the worker's probability of quitting, and increases profits. The third effect is that a lower wage reduces the worker's incentive to remain with the firm, thereby increasing the turnover rate. Combining the second and third effects yield that lowering the wage increases turnover probability. The outside market correctly anticipates the firm's tendency to underpay. That is, a \$1 drop in the pay offered by the current firm reduces the signaled ability (and outside offers) by less than \$1. Thus, the turnover probability is greater as wages are reduced.¹⁷

In this equilibrium, all but the highest ability worker is paid a wage below the wage paid give symmetric information. Therefore worker utility is lower and

¹⁷The result that turnover is higher under full separation may seem a bit at odds with the limit pricing result of Milgrom and Roberts (1982). In their model, full separation under asymmetric information results in the same entry rate as symmetric information. The key difference has to do with the actions available after full revelation. In their model, after the first period price is chosen and costs are revealed, the game is identical to symmetric information. There is no direct effect of the first period price on the entrant. In my model, the wage chosen by the first period employer has a direct impact on the decision made by the employee. The outside market will offer the same wage as under symmetric information, but the current firm's wage is lower. Therefore the worker will quit with a higher frequency than it would under symmetric information.

turnover is less efficient than under symmetric information.

Now that I have characterized and discussed the unique fully separating equilibrium I will discuss pooling. Pooling exists when a range of ability types are all offered a common wage in equilibrium. The outside wage offer is equal to the signaled ability of a worker offered that wage. Pooling equilibria may alleviate some of the problems caused by wage signals. However as long as multiple wages are offered in equilibrium there is excessive turnover reducing expected worker utility.

Proposition 3 *The average expected second period profits for workers offered the highest wage is bounded above by $x^*[1 - G(x^* - \sigma)]$ which is the expected profit under symmetric information. For any two wages offered in equilibrium $w' < w''$ The average turnover rate for workers offered w' is higher than for those offered w'' .*

Proof. See proof of Proposition 2 ■

As above, the tendency for the firm to signal low ability through low wages causes the outside market's beliefs to be less responsive to lower wages. Therefore for lower wage offers turnover is higher and average second period profitability is lower.

While I have shown that any separation in equilibrium induces negative effects, there may be a fully pooling equilibrium that can induce the same turnover rate and expected second period profits as under symmetric information. If such an equilibrium exists it will be unique. All worker types will be offered the common wage equal to the expected ability minus x^* . All risk associated with ability realization is borne by the firm through the common wage. Turnover rates will

be constant for all abilities as a result of the common inside wage and common outside wage.

I now discuss why although the wages are quite different we get the same turnover rates in this equilibrium as symmetric information. In symmetric information although wages are increasing in ability, the difference between inside and outside offers is always $-x^*$. In the pooling equilibrium discussed above the difference between inside and outside wages will also be $-x^*$. The turnover rate is increasing in the difference between the current and outside wage offers. Therefore turnover rates are the same in both the pooling equilibrium and full information.

Proposition 4 *Let \bar{a} be the mean ability of the distribution F . There exists an equilibrium that induces the same turnover rate for every ability type and expected second period profit as the symmetric information model if and only if,*

$$(a_L - \bar{a} + x^*) [1 - G(x^* - \sigma)] \geq \max_w (a_L - w) [1 - G(a_H - w - \sigma)]. \quad (1.16)$$

Further the wage policy $w_i(a_i) = \bar{a} - x^$ is the unique wage policy for this equilibrium.*

Proof. Direct application of Condition 1. ■

Proposition 4 discusses the uniqueness and existence of the fully pooling equilibrium that achieves the same efficiency levels as under symmetric information. However, unlike full separation, this may not always be an equilibrium. The pooling wage could potentially be too high to satisfy incentive compatibility when the firm observes the lowest ability type. To see this, suppose the pooling wage is higher than the marginal product of the lowest type. If the firm observes the

lowest ability type it would never pay a wage greater than marginal product in the second period. Therefore, this cannot be an equilibrium strategy for the lowest ability type. In fact the condition on the pooling wage is even more strict. The condition states that it can never be more profitable for the firm to pay a lower wage and signal the highest ability type than to pay the common wage.

Corollary 1 *If the lowest ability type a_L is sufficiently below the mean ability¹⁸, then any equilibrium will induce lower expected profits for the firm and higher turnover for some ability types than under symmetric information.*

Proof. Direct application of Propositions 1 and 3. ■

Corollary 1 formalizes when asymmetric information will necessarily harm profits and turnover in the second period. This is again when the spread of the ability types is large enough that there is no pooling equilibrium.

I have shown that asymmetric information over worker ability can create additional turnover inefficiencies compared to symmetric information. Ironically signaling ability increases turnover and lowers second period profits compared to symmetric information. In the next section I show how formal pay systems can improve turnover inefficiencies and can thus be preferred by incoming workers.

1.4 Pay Systems

In this section I discuss how pay systems reduce turnover in the second period.

I have shown that the signaling and turnover aspect of wages can increase the

¹⁸The specific condition is $(a_L - \bar{a} + x^*) [1 - G(x^* - \sigma)] < \max_w (a_L - w) [1 - G(a_H - w - \sigma)]$.

difference between inside and outside wage offers. Pay restrictions may be used to restrict the firm's ability to distort wages downward, thus improving second period profitability.

One possible solution to the wage signaling problem is to commit to make any information about ability public (see Bar-Isaac, Jewitt and Leaver 2008).¹⁹ This would change the firm's second period problem to be exactly the symmetric information case. Although it is not the focus of this paper, it could also be an interesting solution to this problem. However in certain cases, this solution may also not be possible as information on worker ability may not be verifiable.

In the following subsections I examine three different types of pay restrictions that firms often employ. These pay systems can improve efficiency through employee retention. Retention improves as second period wages are higher relative to outside offers. These restrictions are put in place prior to the firm observing the ability of the worker. I show that these pay systems can improve efficiency when wages are a signal of ability. Since the model is competitive all efficiency gains are passed on to the worker at the beginning of the working relationship. Therefore workers would prefer to work for firms that use one of these systems if it improves future profitability and future retention. However these restrictions may harm efficiency when information about worker ability is symmetric. In addition to discussing efficiency and second period profitability of such policies I also discuss testable distinguishing characteristics between the two information assumptions.

¹⁹In their paper the firm deals with a traditional adverse selection setting where private information is valuable to the firm and there is no wage signaling. They look at the decision for firms to make private information about workers public. The firm may choose to commit to give up this information in the future if it helps attract new employees.

1.4.1 Single Wages

Perhaps the most strict and glaring example of a pay system is what I refer to as a single wage system. The most obvious example of this type of system is the US federal government's GS system. Employees are paid a specified wage based on experience and education. Pay increases are largely determined by tenure (see Mace and Yoder 2009).

I now describe how this pay system can be incorporated into this model. The firm chooses a second period wage prior to the worker accepting a job with that firm. After the firm observes worker ability the firm is compelled to pay this wage to all existing employees in the second period. The outside market again observes this wage and makes a wage offer just as in the previous model.

Under wage signaling this pay system mirrors a fully pooling equilibrium from the previous section except with much more flexibility. All workers are paid the same wage regardless of ability. By using this system the firm can guarantee an average expected second period profit of any amount up to and including the maximum of V^* . They may even be able to reduce turnover inefficiencies lower than in the symmetric information case. Even if such a fully pooling equilibrium were possible in the wage signaling model equilibrium selection may be unclear. Having the single wage prescribed beforehand allows the firm to commit to a good equilibrium with certainty.

The single wage pay system can create a fully pooling equilibrium when it would not exist otherwise. I have shown in Corollary 1 that if the ability dispersion is too great there is no fully pulling equilibrium with a turnover rate less than or equal to the turnover rate in the symmetric information case. The limitation that

causes this result is the incentive compatibility condition on the lowest ability type. Under the single wage pay system it is assumed the firm can commit to pay all employees the same wage even if it is not incentive compatible after observing ability. Therefore the firm can credibly commit to a fully pooling equilibrium and achieve higher profits.

Even though profit maximization is possible under this pay system the firm will want to pay more than the profit maximizing wage. I have shown that a profit maximizing wage of expected output minus x^* induces too much turnover from a welfare standpoint. Recall that the market is competitive in the first period and all rents go to the employee. Therefore, the employees would prefer the firm to commit to a wage that induced efficient turnover. This wage would be exactly the average output of the workers.

The benefits of such a system are clear under asymmetric information and wage signals. Just as with all pay systems I discuss in this paper, the value comes from reducing the wage signaling problem. However, when worker ability is common knowledge single wages can severely harm second period profits and alter turnover through what I call reverse adverse selection. I now discuss the effect of such a system when information about worker ability is symmetric.

Proposition 5 *If the firm commits to a single wage in the second period (\bar{w}) and information on ability is symmetric, then turnover rates increase with ability. Also the average ability of a worker who stays is less than of a worker who moves.*

Proof. In appendix. ■

Proposition 5 shows that when the firm pays a single wage to all of its employees the outside market lures the more productive workers away with a higher wage.

Lower ability workers have less compelling outside options and are more likely to stay with the firm. As a result, the average ability of quitters is higher than average ability for those who remain. This results in a sort of reverse adverse selection where turnover happens disproportionately for the more able workers.

This case generates a different prediction than adverse selection models tested by Gibbons and Katz (1991) and others.²⁰ Symmetric information and single wages predicts that workers displaced after a plant closing will have lower ability than those who quit beforehand. Single wages under the wage signaling model predicts no relationship between turnover probability and ability. Therefore the average ability of the stayer and the quitter will be the same. The average ability of a quitter and a worker who leaves because of a plant closing should be the same.

I have just discussed the value of a strict single wage system in overcoming the wage signaling problem. However, in many instances, it may be impossible to commit to this policy. If it is learned that some of the workers have a marginal product below this wage the firm has an incentive to let that worker go. In the single wage system, while average ability is above the single wage some ability types are paid more than their marginal product. Once a worker is revealed to have low ability, a firm may want to fire the worker rather than keep the low ability worker at the fixed wage. The feasibility of the single wage system rests on the commitment ability of the firm to not fire. Unlike many firms the US Federal Government may be able to maintain rigid enough bureaucracies to apply this type of system.

In addition to the extreme rigidity of this system there is another reason why this may be impractical. Although worker effort is not modeled in this paper,

²⁰Recall these adverse selection models predict that the ability of the workers who are laid off is lower on average than ability for those who are terminated due to a plant closing.

a single wage system would eliminate any incentives for a worker to exert effort. Just as in the fully pooling equilibrium discussed above, this single wage system may induce perverse incentives to not work. The following pay systems still at least partially negate the wage signaling problem while maintaining some wage dispersion.

1.4.2 Pay Scales

Another common type of pay system is pay scales. These are ranges of pay that the firm must stay within when offering compensation to employees. The ranges are usually tied to observables such as job assignment, firm tenure and/or experience.

As previously discussed the signaling role of wages reduces efficiency through high turnover. Since the fully separating equilibrium is unique any pay scale that is narrower than this equilibrium wage profile must result in some pooling. I now show how such pooling can improve welfare and can thus be preferred by potential employees.

The model is identical to the one above except that the firm can now commit to pay wages within a range $[\hat{w}_l, \hat{w}_h]$. The pay scale is chosen before period 1. Once firms have chosen their pay scales and observed worker ability they play the game just as in Section 1.3.2. I first examine the effect of pay scales when wages signal ability. As before there is no unique partially pooling equilibrium to this game. I narrow the new set of equilibria according to the following conditions. These rules will uniquely define the equilibrium wage profile given the pay scale.

Definition 2 *Given a pay scale $[\hat{w}_l, \hat{w}_h]$ the constrained separating equilibrium is*

an equilibrium with the following characteristics.

- 1) Any pooling happens at the extreme wages.*
- 2) The average expected second period profit on the highest paid worker is at its maximum.*

The first condition guarantees that this equilibrium somewhat resembles the outcome of symmetric information.²¹ By restricting the equilibria where any pooling occurs at the boundary I generate additional testable predictions distinguishing wage signaling from symmetric information. The second condition is naturally satisfied by a fully separating equilibrium. This condition narrows the equilibria to those more comparable to the fully separating equilibria. These two conditions guarantee uniqueness of this type of equilibrium. Also if the upper wage boundary is high enough and the lower wage boundary is low enough this equilibrium is exactly the fully separating equilibrium.

The constrained separating equilibrium is determined from the top down. The maximum wage offer determines the abilities pooled at the top and the separating wages for the middle range. The minimum wage offer determines the ability types that are offered the lowest wage.

In showing the benefit of pay scales I look at restrictions from the top and the bottom separately. Starting from the fully separating equilibrium (i.e. $\hat{w}_h \geq w_i(a_H)$ and $\hat{w}_l \leq w_i(a_L)$), I now show that reducing the maximal wage below the highest fully separating wage ($w_i(a_H)$) improves profits.

²¹As in the earlier wage signaling model there may be partially pooling equilibria where multiple ability types are offered the same wage within the extreme wages. This internal clumping never occurs when information on worker ability is symmetric. Therefore empirical evidence of such clumping would support the wage signaling model but not the symmetric information model. However observing no clumping within the pay scale would be inconclusive.

Proposition 6 *Compared to the fully separating equilibrium, there exists a range of restrictive upper pay bounds (less than $w_i(a_H)$) that improve efficiency through reduced turnover.*

Proof. In appendix. ■

The proof of Proposition 6 shows that turnover rates can be improved by upper wage restrictions in two ways. First, reducing the maximal wage causes more workers at the top of the ability distribution to have the second period profit maximizing turnover rate. This lowest turnover is now expected on a larger portion of the workforce. Second, increased pooling at the top reduces turnover rates for all worker types through the individual compatibility restrictions. For some small amount of pooling these constraints become more relaxed and the equilibrium wages are higher for all workers. While some amount of pooling at the top can improve this, if the upper pay restriction is too low, wage offers on some ability types may decrease. In this case, the workers who are not pooled may have higher turnover than under full separation. That is, the first effect is always welfare improving while the second effect increases welfare for at least some upper pay restrictions. Therefore there will always be some upper pay restriction that improves efficiency through lower turnover.

Now that I have shown that an upper pay boundary can improve profits I examine the role of the lower pay restriction. The result in this case is not quite as strong but it illustrates that such lower boundaries may be useful in some cases.

Proposition 7 *Given any upper pay restriction (\hat{w}_h) in a constrained separating equilibrium, a lower pay restriction improves efficiency if the distribution of ability types receiving that wage has low enough variance.*

Proof. In appendix. ■

The efficiency increases of Proposition 7 come from increasing the wages of the lowest ability types. The separating equilibrium results in particularly high turnover when the firm observes a low ability worker. This high turnover is caused by low equilibrium wages relative to outside offers. A lower bound on wages can improve this problem through a commitment to pay higher wages on average. For the average worker within this group, this wage will increase his wage and reduce turnover inefficiencies. However this pooled wage will also cause the highest ability workers within that range to turnover more than they would otherwise. Because there are both winners and losers, the wage restriction has an ambiguous effect on welfare without knowing more about the distribution of ability. We know that the separating wage profile is convex meaning that wages increase with ability at a faster rate for higher ability workers. Decreases in efficiency in high ability workers from the pooled wage can affect total inefficiencies more than the turnover reduction for low ability workers. However if the ability distribution is sufficiently tight about the mean within this group, then such wage restrictions will improve welfare.

Pay Scales with Symmetric information.

Now I examine the effect of pay scales when worker ability is observed by the outside market. In this case, I show that the current firm no longer has the tendency to distort the wage offer away from profit maximization. Any binding pay restriction such as pay scales results in lower second period profits. Evidence of symmetric information when there are pay scales is found through the relationship between turnover rates and wages.

Proposition 8 *If there is symmetric information about worker ability, any upper wage $\hat{w}_h < a_h - x^*$ results in a range of abilities being offered that wage $[\hat{w}_h + x^*, a_H]$. Further, average turnover rates are higher and expected second period profits are lower on such workers. Similarly, any lower wage $\hat{w}_l > a_L - x^*$ results in a range of abilities being offered that wage $[a_L, \hat{w}_l + x^*]$. Average turnover rates and expected second period profits are lower on such workers.*

Proposition 8 shows the effect of a binding restriction in the absence of wage signaling and turnover. Just as traditional labor theory suggests, such restrictions change the turnover rates for the workers whose wages are changed under these restrictions. If the unconstrained optimal wage is not available under the pay scale, the firm chooses the closest feasible offer. The highest (lowest) ability workers are paid the upper (lower) wage boundary. Compared to profit maximizing, the highest (lowest) ability workers are paid too little (much) and quit more (less) frequently resulting in lower profits. Since turnover rates are already inefficiently high, it is possible that lower pay restrictions may improve welfare. Again, since all rents are given to the worker, such restrictions may be preferred by incoming employees. Upper pay restrictions further reduce welfare and are not preferred by the workers.

Now that I have determined the characteristics of pay scales under both cases I now examine the distinguishing features of the two models. Without looking at the turnover rates the wage profile for the symmetric and the asymmetric can look quite similar. In both cases, restrictive pay scales result in pooling at the extreme wages. Determining which of these models best describe the real world is done though examining turnover rates at and within the extreme wages.

The wage signaling model predicts turnover rates decrease with the wage whether or not there is clumping at the extreme wages. That is, turnover rates are highest for the lowest paid workers and lowest for the highest paid workers. Also turnover rates are negatively correlated with wages for workers in between the extreme wages.

The turnover predictions for the symmetric information model are quite different. The symmetric information model predicts that turnover rates are uncorrelated with the wage if there is no clumping at the boundary. If there is clumping at the highest wage, then turnover rates at the high wage are higher than for other workers. If there is clumping at the lowest wage, the turnover rates at the low wage are lower than for the other workers. Therefore the two models generate vastly different predictions on how pay scales affect turnover rates at the high and low wages.

The limited empirical work on pay systems seems to show that upper pay bounds do not adversely affect turnover. Gibbs and Hendricks' (2004) analyze a single firm that used a pay scale system. They find evidence of clumping at the high end of the wage scale while no evidence of increased turnover at the maximal wage offer. This finding is consistent with the asymmetric information and wage signals, but not with symmetric information.

1.4.3 Budgets

Another common pay restriction is the use of budgets (see Merchant and Manzoni 1989, and Hansen and Van der Stede 2004). Here the governing body of the organization sets expenditure restrictions on lower managers. The lower managers

then have the discretion to allocate the funds among employees.²² Examining the role of budgets is a huge topic within the field of accounting (ex. Demski and Fetham 1978, and Shih 1998). This literature is largely based on principle-agent problems between the firm and lower level managers. The underlying theme of this previous work is that budgets allow the firm more control over the internal workings of the firm. The motivation for using budgets in this paper is very different and complementary to the vast work in the accounting literature. In this paper, I argue that firms may use budgets as a commitment device against itself rather than to induce optimal behavior from an agent. I then show that these same types of budgets harm profitability when information on worker ability is common knowledge.

I start with the same assumptions as in the simple wage signal model (Section 1.3.2) with a few exceptions. First, I introduce managers into this game. Each new employee i is assigned to a manager in the first period. For each set of new employees a manager is given $n \geq 1$ new workers with abilities drawn i.i.d. from distribution F . After the first period of work the manager observes ability and chooses wages for the workers under his span of control for the second period. The objective for a manager is slightly different than the firm's. Instead of choosing wages to maximize profit, the manager is maximizing total expected output subject to a budget constraint. This budget constraint is set by the firm. In this model I have set the budget constraint as restricting the maximum expected combined wages of the group under the manager's span of control. The manager must set wages so that on average, the combined wages must fall below some maximum.²³

²²There may be other reasons not modeled in this paper why a firm would want to use budgeting. Even so, this paper generates testable implications of the effect of the budget restriction on turnover rates.

²³An alternative assumption might be a restriction on total wages offered rather than expected wages accepted by workers. This restriction does not change the results of the wage signaling model, but it can have some effect under symmetric information. The restriction as presented in

That is after observing the draw of abilities a_1, \dots, a_n the manager chooses wages to solve,

$$\left\{ \begin{array}{l} \max_{w_1, \dots, w_n} \quad \Sigma_{i=1}^n a_i [1 - G(b_i(w_1, \dots, w_n) - w_i - \sigma)] \\ \text{st} \quad \Sigma_{i=1}^n w_i [1 - G(b_i(w_1, \dots, w_n) - w_i - \sigma)] \leq K \end{array} \right\}. \quad (1.17)$$

Again, consistency requires that the beliefs associated with a wage offer $b_i(\cdot)$ must be the expected ability of a type receiving that wage offer. I also assume that if the wage offers of any two employees are switched then the beliefs about their productivity switches as well.²⁴

One option for the firm setting the budget restriction is to set the budget to maximize expected second period profits. Under wage signaling it is possible to get an equilibrium that is very similar to full information. This is done by setting the wage equal to the expected wage bill of the group if information was symmetric.

Proposition 9 *If the firm sets a budget $K = n(\bar{a} - x^*)$, then there exists equilibria that induce the same expected profits as under symmetric information. However there may be other equilibria that do not.*

Proposition 9 shows us that using budgets allow for equilibria where the firm achieves maximal profits in the second period. This is because the person setting the wage no longer has the incentive to reduce costs below the budget restriction.

the paper is the easier to interpret and generates cleaner and more realistic predictions.

²⁴This guarantees that in equilibrium wages will be increasing with ability. To see that this must be true assume that the opposite is possible. If given two ability types a higher ability worker also has a higher turnover rate in equilibrium, the managers can switch the employees' wages and gain more expected profits.

In staying within budget, the manager is concerned with the signaling role of the wage offer. But now there exists equilibria where any equilibrium wage offer induces the same turnover probability. Therefore the manager is indifferent between wage offers as long as he stays within budget.

The equilibria that induce maximal expected profits are not unique. However, they have some important common characteristics. In any such equilibria, the turnover probability is the optimal $G(x^* - \sigma)$ for all possible equilibrium wages. This imposes a restriction on the wage policy. For any wage offered in equilibrium, the expected ability of a worker offered that wage must be x^* higher than that wage. This is now possible in equilibrium since turnover rates are the same regardless of the wage offer. In this equilibrium, the manager has no incentive to distort the wage offer. The wage profile must set wages in a way that expends the entire budget while signaling ability. One possible equilibrium is the perfect pooling equilibrium of all workers offered $\bar{a} - x^*$. In this equilibrium the expected ability of any worker receiving that wage would be the mean of the ability distribution. Another profit maximizing equilibrium is for each worker to be paid x^* less than his i th order statistic. In this equilibrium the manager observes ability and orders workers on ability from 1 to n . Then the manager pays each worker i x^* less than the expected ability i th order statistic from a random draw of n . As a result the outside belief of the i th highest wage offer is exactly the expected i th order statistic. Both of these wage profiles are an equilibrium that induces maximal expected profit.

Even though second period profit maximization is possible workers would prefer the firm to set the budget even higher to maximize efficiency. This budget restriction would be the expected combined marginal product of the group. The

equilibrium would work the same as above only without subtracting x^* .

The use of budgets is an example of how centralizing to some extent the wage setting decision can improve profitability. By taking the incentive to underpay employees away from the decision maker, the firm is able avoid excessive turnover.

One major criticism of strict formal pay systems is that such systems lack the flexibility necessary to react to the outside market. Under asymmetric information and wage signals I have shown such rigidity makes pay systems valuable. However, as I show next, under symmetric worker information rigid budgets reduce profitability no matter how well executed. This lack of flexibility causes workers to be either over or under paid compared to the optimal wage.

I now analyze this model under the assumption of symmetric information on worker ability. All other assumptions are the same. Since the wage no longer acts as a signal the problem now becomes,

$$\left\{ \begin{array}{l} \max_{w_1, \dots, w_n} \quad \Sigma_{i=1}^n a_i [1 - G(a_i - w_i - \sigma)] \\ st \quad \Sigma_{i=1}^n w_i [1 - G(a_i - w_i - \sigma)] \leq K \end{array} \right\}. \quad (1.18)$$

The solution to this problem will depend on the realized ability level of the n workers assigned to the manager as well as the budget constraint.

Proposition 10 *If ability is common knowledge strict budgets reduce the average expected second period profit on workers.*

Proof. In Appendix ■

Proposition 10 shows that under symmetric information any such budget restriction lowers expected profits. This result comes from the assumption that the

budget choice is made prior to the realization of abilities. A budget that is not determined perfectly results in either too much turnover or wages that are too high. Either effect reduces expected firm profits. It is possible for the budget to be the perfect size so that expected profits are at their maximum. This would depend on the realization of abilities of the workers. Because budgets are set before observing ability, the probability that the firm sets the budget perfectly is zero. Any attempt to set budget restrictions results in lower expected profits for the firm.

I have shown how the inflexibility of the budgets can harm firm profits. I now examine how such budget restrictions affect worker turnover and compensation. This provides a testable implication on the effect of budgets on worker turnover.

Corollary 2 *Turnover probabilities of all workers in a unit increase when the average ability of workers within that unit increases for any budget restriction. Also, wages and turnover rates for higher ability workers are affected more by the difference in the sum of unrestricted wages and the budget restriction.*

Proof. In Appendix ■

Corollary 2 explains the effect on wages when the actual budget does not match the realized abilities of the group members. When the budget is too high then wages for all workers increase and turnover decreases. When wages are too low, then all workers take a pay cut and turnover for all workers is lower. It also states that the effect is greater for higher ability workers.

The intuition behind this result is as follows. If the budget is high compared to combined worker ability, the manager's main concern is improving expected output

through retention. As a result the manager increases his retention efforts on all workers. He is particularly worried about retaining his most productive workers. If the budget is low, the manager is primarily concerned with wage reduction. All workers receive lower wages. However, the highest ability workers are also the highest paid. Lowering the highest wages and increasing turnover rates has the greatest effect on reducing expected wage costs. Therefore the higher ability workers are affected more by the budget than the lower ability workers.

Although budget restrictions will always reduce second period profits they may improve efficiency if set high enough. The unrestricted full information wage profile still induces inefficiently high turnover. This inefficiency may be reduced if the budget causes wages to be closer to marginal product. However maximum efficiency is not attainable due to the randomness in realized worker ability.²⁵

As noted above the effect of budgets on turnover rates have different predictions depending on the information structure of worker ability. I now suggest tests that can distinguish a symmetric information result from wage signaling. Under wage signaling, budgets can cause turnover probabilities to be the same across ability types. Therefore one worker quitting has no predictive power on other quits within that group. In symmetric information, a quit signals that the budget is too small for the ability stock. Therefore quits will be correlated for workers within the span of a manager. Another prediction of this model is that the quit rates of higher ability workers are affected more by the relative budget size compared to ability stock. This means that quitting tendencies are more strongly correlated for higher paid workers.

²⁵Maximum efficiency only occurs if the inside and outside wages are the same. With symmetric information, the outside wage is always equal to marginal product. But, since the budget is determined before the realization of ability for the ability of the group the budget will always be either too high or too low to induce the efficient turnover rate.

1.5 Discussion

Information plays a crucial role in this turnover model. I have examined the turnover model separately for wages as signals of ability and under symmetric information. There is another possible information assumption that I now discuss. That assumption is that neither the current firm wage offer nor worker ability are observed by the outside firms.

In this case the wage offer has no signaling role. All workers are offered the same outside wage offer regardless of their ability or current firm wage offer. In this case, the inside firm offers a higher wage to its more productive employees to reduce turnover on its highest producers. Since turnover is higher for lower ability workers the average ability of workers who quit is lower than for workers who do not. The single outside wage offer reflects the lower expected ability of the worker who quits. Since workers on average expect to receive a lower wage in the outside market, few workers move firms. This model is closely related to the existing literature on adverse selection in the labor market. Therefore I refer to this model as adverse selection.

Depending on the observability of ability and wage offers there are three models. They are symmetric information, wage signals, and adverse selection. While there are similarities, I now suggest some empirical tests to distinguish between these three models in the absence of formal pay systems. The first test is whether wages are correlated with turnover probability. The symmetric information model is the only one of these models where wages are not correlated with turnover rates. In the other two models turnover and the wage offer are negatively related. The second test is whether pre quit wage offers are correlated with post quit offers. In both the symmetric information and wage signaling model pre quit wage offers

predict outside firm wage offers. For adverse selection, wages and productivity at the current firm are unobserved by the market and do not predict outside wage offers.

In summary the correlation between turnover rates and wages give evidence on the information structure of the labor market. In the absence of formal pay systems, if wages are uncorrelated with turnover this suggests symmetric information. Current firm wages that are correlated with turnover probability as well as outside offers support wage signaling. If current firm wage offers are correlated with turnover but not correlated with outside wage offers this suggests adverse selection.

1.6 Conclusion

In this paper I attempt to explain a possible reason for why firms use formal pay restrictions. I argue that formal pay systems can help the firm to become more attractive to potential employees. Pay systems allow the firm to commit to wage profiles in order to reduce the losses associated with excessive turnover. I have also shown how these pay systems may be harmful or helpful under different information assumptions. In addition I have generated various empirical predictions depending on the information structure and the various pay systems considered.

I argue that reducing excess employee turnover is an important new avenue for welfare improvement. If these pay systems indeed reduce inefficient levels of turnover, it is important for employers to understand such contracts. This wage policy would allow the labor market to respond to market forces yet fully benefit from individual workers tastes. While I have discussed three distinct types, a

more complete description and analysis of all possible wage systems would likely provide further insight.

CHAPTER 2

TURNOVER AND PROMOTION

2.1 Introduction

Employee turnover has been of great importance to the business world and economists for some time. A vast theoretical and empirical literature has been developed to understand labor mobility. Similarly, the last few decades have seen an increased effort to try to understand the issue of job assignment. While employee promotion and turnover have been studied separately, this paper is among the first to examine the relationship between these issues. A number of empirical studies suggest a negative relationship between turnover and promotion (see Carson et al. 1994; Saporta and Fajourn 2003). To our knowledge, no theoretical work has examined this relationship. Using a simple, intuitive, yet largely unexplored turnover process, we explain this and other established stylized facts relating to job assignment.

Our turnover mechanism also explains what has been an important disconnect between economic theory and business practice. Employee turnover is a serious matter for firms. Firms often hire consulting groups to help manage employee retention or the “cost of employee turnover.” Explicitly these turnover costs include search costs, training costs, setup costs, etc. These costs can be substantial even for the lowest paid worker.¹ While turnover has been examined in many contexts, current economic theories have difficulty justifying the heed given to employee retention.² Our turnover process provides a theoretical basis to the idea of the

¹For example, in their study of four different hotels Hinkin and Tracy (2000) estimate the turnover cost of a front-desk associate to be from 5 to 12 thousand dollars. Over half of this figure is non-explicit “productivity” costs. See also Cascio (1999) and Wasmuth and Davis (1983) for additional examples of costs of employee turnover.

²Under traditional theories, when counteroffers are possible, the firm will never allow a worker

“cost of employee turnover.”

The source of turnover in our model is (firm and time specific) private random utility shocks experienced by the worker.³ The current employer is able to exploit these shocks through wage offers that are below the employee’s marginal product. Therefore, the employer captures rent from the worker as long as the worker ‘enjoys’ working at the firm. If the worker receives a shock that is low enough, it will quit the firm and work elsewhere. When workers leave the current firm, this creates a loss in current and future expected profits. We interpret the loss in current and future expected profits as the cost of turnover.

Another (somewhat less important, yet instructive) element of this model is switching costs. These represent moving expenses, emotional stress or other costs associated with changing employers. These costs allow for a more realistic level of turnover as well as a way to examine how different turnover rates affect the promotion decision.

In order to isolate the effects of these mechanisms we first build an infinite-horizon model with only one job assignment. We assume that workers are equally productive at all firms and that this productivity is common knowledge. After a period of work the worker observes his taste for the following year. Since the worker only has information about the current firm, the current firm can capture rent from

to leave before offering him a wage equal to his marginal product. If workers are paid their marginal product, turnover of a single worker will not affect the firm’s profitability. Under some restrictive assumptions, existing theory can generate endogenous turnover while firms pay a wage below marginal product. For example, a model with random outside wage offers that assumes away counteroffers may have this feature (see, Munasinghe and O’Flaherty 2005). However, assuming away counteroffers does not seem to be the a realistic approach to this question.

³These shocks have a natural interpretation. They are the various non-pay aspects of a job whose importance to the worker is unknown to the firm. We assume that these shocks change over time. Workers can only predict future shocks for their current employer one period ahead. Since the worker has information about the current employer, the current employer is able to exert a kind of monopsonistic power over the employee.

the worker's private information. This is done by a lower wage offer. Each period, the current firm offers a wage below marginal product. If the realized taste shock is high enough to compensate for the lower wage offer the worker remains with the current firm. If the realized taste shock is too low, it switches employers. Once the employee leaves his current firm, that firm loses claim to all current and future profits from that employee.

We see from this simple model that private worker tastes generate a profit stream for the current employer. Because current firms are able to capture future rents, firms will be willing to pay up front incentives to establish a relationship with new employees. This explains large wage increases associated with switching jobs as well as signing bonuses and other up-front payments made at the beginning of the match.

We later extend this model into a simplified Gibbons and Waldman (1999, 2006) framework of job assignment.⁴ The worker starts his life more productive at the low level job. Over time, the worker becomes increasingly more productive at the higher job. We assume that there is a one-time general training or "learning-by-doing" cost associated with promotion (see Brown 1989, Gathmann and Schonberg 2007, Shaw and Lazear 2007, and Dohmen 2004 for empirical evidence consistent with this assumption). If the current firm does the training, the worker has the ability to change employers immediately after the firm pays the training costs. If the current firm doesn't promote, the outside firms can offer promotion to attract the potential employee.

In equilibrium, an employer wants to maintain its relationship with its current

⁴In order to isolate the promotion, turnover and wage dynamics we abstract away from the learning elements of the Gibbons and Waldman framework. We assume that worker productivity is deterministic and common knowledge.

workers to extract the aforementioned rents over time. However, the firm is also concerned with paying the up-front cost of training without a guarantee of the worker staying with the firm. As a result, the firm will delay promotion past the point that it is first-best optimal. The outside firms will start offering promotions before the inside firm. Because the worker becomes increasingly more productive at the higher job, the outside wage offers improve relative to the current-firm wage offer. The worker's turnover probability increases during this time. Eventually the strain of the outside market's wage offer induces the inside firm to also offer promotion (and a large pay increase) to its loyal employee. Once the worker is offered promotion (either from the current firm or not) his subsequent turnover probabilities fall to their lowest level. Therefore, we suggest that the decreased turnover at promotion comes from an alleviation (of outside market pressures) that happens at promotion.

In addition to explaining the negative relationship between turnover and promotion we generate a large set of well established stylized facts regarding turnover, promotion and wage dynamics. In setting up our model, we assume away many traditional explanations used to explain these various established stylized facts (such as, asymmetric information on ability types, learning, or contracts). We do not claim these other forces are unimportant. However, we claim that this new class of labor market models can explain new stylized facts while being consistent with established patterns in the data.

The remainder of the paper is organized as follows. Section 2 reviews the current literature that relates to this paper. Section 3 presents a simple model that introduces the turnover mechanism. Section 4 examines this model with job assignment. In section 5 we outline an extension that incorporates a continual ele-

ment of the worker's taste for the firm. We discuss the implications and alternative explanations in section 6. Section 7 concludes the paper.

2.2 Literature review

The last three decades have produced a great deal of research designed to help understand the relationship between the worker and the firm. This paper is among a small set of personnel economics models that take important earlier models⁵ (ex. Becker 1962, Lazear and Rosen 1981, Rosen 1982, Greenwald 1986, Gibbons and Waldman 1999) and incorporate private worker taste (or disutility) shocks. These new models use this turnover mechanism to generate more realistic wage and turnover predictions (ex. Acemoglu and Pischke 1998, Barron and Berger 2006, Schonberg 2007, Ghosh 2007). I also use this mechanism in the other two chapters of this dissertation. In both chapters 1 and 3 I examine firm policies and their relation to turnover efficiency. In Chapter 1, I discuss formal pay systems. In chapter 3, I consider employee bonding.

The work of Ghosh (2007) is the paper most similar to this chapter. He adds the turnover mechanism to a three period Gibbons and Waldman (1999, 2006) framework. His model naturally produces many of the same results as Gibbons and Waldman. In addition he shows that firm specific human capital accumulation creates a negative relationship between firm tenure and turnover.

Although built on the same framework, our paper differs from Ghosh in a number of important ways. First, and most importantly, his paper does not examine

⁵Another notable work that is less relevant to this paper is Harris and Holmstrom's (1982) discussion on insurance.

the relationship between turnover and promotion. Our results largely stem from our assumption of job (or task) specific training costs. Further discussion will be included later in the paper. Second, we abstract away from the learning elements of the Gibbons and Waldman framework. Third we use an infinite horizon model. This feature is absent in all the aforementioned ‘disutility’ shock based turnover modes. Assuming an infinite-horizon helps to disentangle the effects of promotion, tenure, learning, etc. on wage dynamics from ‘end of life’ effects.⁶

One of our key predictions in this paper is the negative relationship between turnover and promotion. In the field of psychology a large empirical literature argues that ‘promotion satisfaction’ or ‘promotion opportunities’ reduce turnover rates (ex. Cotton and Tuttle 1986). However Carson et al. (1994) use a meta-analysis on numerous such studies and show that actual promotion rather than ‘promotion satisfaction’ or ‘promotion opportunities’ reduces turnover. Similarly, Saporta and Fajourn (2003) used longitudinal data on a single firm and found a similar result. They found that the number of promotions reduced turnover rates for both professional and managerial workers.

In addition to predictions on turnover we also generate a well established relationship between wage and promotion. Particularly, compensation is largely associated with job assignment (ex., Baker, Gibbs and Holmstrom 1994).

This is not the first theoretical work to generate this prediction. Early theoretical work assumes that jobs can be ordered by their responsiveness to ability (see, Rosen 1982, Waldman 1984b). Jobs in which productivity depends more on

⁶Firms will pay more to keep (or hire) an employee if that employee will provide future rents. The worker will remain with the current firm for a lower wage if future expected utility is higher (relative to outside option). That is, future profit raises the current wage offer. Future worker utility lowers the current wage offer. In a finite horizon model, the concern for the future is not the same in all periods. The wage decision will be different depending on the number of periods remaining. We avoid these complications by using an infinite horizon model.

ability will be filled by more able workers. The wage distribution becomes more skewed than the ability distribution due to comparative advantages (see, Sattinger 1975). The ordering of jobs alone does not sufficiently explain the wage premium at promotion. Asymmetric information, (e.g., Waldman 1984a, Bernhardt 1995) learning, (see, Gibbons and Waldman 1999) and tournaments (see, Lazear and Rosen 1981) all provide valuable insights. However, none of them address non-degenerate turnover.

In our paper, turnover is a key element and is modeled in a novel way. Although less relevant to this paper there is a vast literature that discusses other types of turnover mechanisms. These papers on job search and matching give us valuable insights into the turnover process (ex. Jovanovic 1979a, 1979b; Burdett 1978). In these models, turnover is a result of economic agents attempting to improve expected match quality. Imperfect information and learning gradually change worker and firm matches over time.

While these models are able to generate a number of empirical predictions they do not discuss why firms are concerned with attracting and keeping individual employees. When it is assumed that workers are paid their marginal product (ex. Jovanovic 1979a, 1979b) there is no real loss to the firm when the worker leaves. Another popular assumption is that workers are given a fixed proportion of their match specific rent (ex. Mortensen 1978). Under this assumption, firms do incur an economic loss when the worker switches employers. However the assumption of a fixed sharing rule of the match specific rent is often not realistic. In many real world settings (where wages are not fixed), a firm would give up additional rents to retain a worker. In models of this sort that allow for counter-offers, the ‘cost of employee turnover’ is still unexplained.

Another key element in this paper is the required training at a change in job assignment. In addition to formal training, this may also be interpreted as learning-by-doing. The task specific training or learning-by-doing is similar to Gibbons and Waldman's (2004, 2006) discussion on task specific human capital.⁷ The learning-by-doing interpretation is compelling as there is ample empirical evidence to support it in many cases. (see Brown 1989, Dohmen 2004, Gathmann and Schonberg 2007 and Shaw and Lazear 2007)

Although training is task specific in this model, we assume that it is transferable between firms. The seminal work by Becker (1962) argues that the costs and rewards of general skills training are entirely born by the worker himself. Recent empirical studies (ex. Loewenstein and Spletzer 1998, and Manchester 2008) show that firms do provide non trivial general skills training. In fact Loewenstein and Spletzer (1998) find that firm sponsored general training increases wages. While Manchester (2008) shows that general training lowers turnover rates.

A number of works try to explain why firms may in fact offer such training (see Katz and Ziderman 1990, Chang and Wang 1996, and Acemoglu and Pischke 1998 and 1999). Katz and Ziderman (1990), and Chang and Wang (1996) look at asymmetric information in relation to general training.⁸ Acemoglu and Pischke (1998 and 1999) also discuss the asymmetric information friction as well as labor unions and complementarities with firm specific human capital. They argue that such frictions 'compress' the labor market wage profile compared to the distribution of marginal product. This compression allows firms the ability to capture some of the returns to training. Our model lacks these types of labor market frictions. We

⁷Other theoretical models of learning-by-doing include Arrow (1962) and Hirsch (1956). Both models focus on the investment aspects of learning-by-doing and not the promotion decision.

⁸Chang and Wang (1996) argue that asymmetric info on training levels will lower actual general skills training levels. Katz and Ziderman (1990) argue that such asymmetries increase training.

suggest an additional complimentary reason why the firm engages in such practices.

2.3 The Turnover Mechanism

In this section we introduce the turnover mechanism that drives our later results. We develop intuition on the source of turnover costs and the rents associated with attracting and retaining individual workers. We assume that worker productivity is full information and place little structure on its dynamics in this preliminary model. In the following section, we add structure to productivity dynamics as we examine how turnover relates to job assignment.

2.3.1 Simple Model

There is free entry into production. All firms are identical. The only input is labor and it is inelastically supplied each period. A worker's career lasts for infinite periods. All firms and workers have a common discount factor of $\delta \in (0, 1)$. Worker i 's marginal product at time t is denoted as $a_{i,t} \in \mathbb{R}_{++}$. All information about productivity is common knowledge.

Worker i 's utility consists of current wages, a taste shock and a switching cost $\sigma \geq 0$ (if incurred). The switching cost is incurred only when the worker switches employers. Worker i 's taste shock for firm k at time t is denoted as $\lambda_{i,t}^k$ and is private information for the worker. All λ 's are drawn independently from a commonly known continuously differentiable distribution G . The distribution G has a zero mean and a continuously differentiable density g . The hazard rate $H(\cdot) = \frac{g(\cdot)}{1-G(\cdot)}$ is non-decreasing on its support. To ensure that turnover is sometimes efficient,

we restrict the infimum of the support of G to be $< -\sigma$. The worker cannot predict his taste shock for any future period. Further, the only period t taste shock worker i knows is for his current firm.⁹ Throughout this paper we refer to worker i 's current firm at time t (denoted as $k_{i,t}$) as the firm where worker i was employed at time $t - 1$. The outside firms for worker i at time t are all other firms except the current firm. We denote $U_{i,t+1}^k$ as worker i 's expected future utility if he works for firm k in period t . If worker i accepts a wage $w_{i,t}^{k_{i,t}}$ from his current employer at time t , his utility is,

$$u_t = w_{i,t}^{k_{i,t}} + \lambda_{i,t}^{k_{i,t}} + \delta U_{i,t+1}^{k_{i,t}}. \quad (2.1)$$

The worker incurs the same switching costs for all outside firms. If he decides to leave the current firm, and accepts a wage $w_{i,t}^{\bar{k}}$ with firm $\bar{k} \neq k_{i,t}$ his expected utility at t is,

$$u_t = w_{i,t}^{\bar{k}} - \sigma + \delta U_{i,t+1}^{\bar{k}}. \quad (2.2)$$

If worker i accepts a wage offer from an outside firm, he will choose the firm that maximizes this expression. We assume that, if indifferent the worker remains with the current firm. Therefore, the worker will stay with his current employer if,

$$w_{i,t}^{k_{i,t}} + \lambda_{i,t}^{k_{i,t}} + \delta U_{i,t+1}^{k_{i,t}} \geq w_{i,t}^{\bar{k}} - \sigma + \delta U_{i,t+1}^{\bar{k}} \quad \text{for all } \bar{k} \neq k_{i,t} \quad (2.3)$$

Recall that the worker only has knowledge of his taste shock for the firm that he worked for in the previous period ($t - 1$). The worker only knows this value for the current period (t) and not beyond. Therefore the only difference in expected utility between the inside and outside options will occur in the current period i.e.

⁹We have not yet explicitly defined the outside option for the worker. It seems natural that wages should be greater than zero. In order to avoid uninteresting corner solutions or negative wages we impose the restriction that for each period $a_{i,t} > x^*$ (where x^* uniquely solves $x^* \frac{g(x^* - \sigma)}{1 - G(x^* - \sigma)} = 1$). However, if we instead allow negative wages and assume that there is no outside option the results of the model are unchanged.

$U_{i,t+1}^{\bar{k}} = U_{i,t+1}^{k_{i,t}}$ for all t and \bar{k} . If worker i enters the outside market, he will accept the highest outside firm wage offer $\bar{w}_{i,t}$.

The employment decision of worker i at t reduces substantially. The worker will remain with his current employer if,

$$w_{i,t}^{k_{i,t}} + \lambda_{i,t}^{k_{i,t}} \geq \bar{w}_{i,t} - \sigma.$$

The worker will leave the current firm with probability,

$$\Pr(w_{i,t}^{k_{i,t}} + \lambda_{i,t}^{k_{i,t}} \geq \bar{w}_{i,t} - \sigma) = 1 - G(\bar{w}_{i,t} - w_{i,t}^{k_{i,t}} - \sigma). \quad (2.4)$$

The payoff for the firm is simply the marginal product of its employees minus the wage costs. A firm's wage decision for a given worker i , depends on whether the firm is worker i 's current firm or not. Since firms are identical, we restrict our analysis to symmetric equilibria. To ease notation we drop the superscript $k_{i,t}$.¹⁰

At the beginning of each period, the current firm gives a wage offer to worker i . The worker decides to accept the offer or enter the outside market. Bertrand like competition in the outside market forces zero expected profits for the winning outside firm. Therefore the highest wage offer ($\bar{w}_{i,t}$) will be equal to the current marginal product of the worker plus the expected discounted rent of being the current firm in period $t + 1$.

The definition of equilibrium that we use is symmetric Markov Perfect Nash Equilibria. In this model it means that a firm's actions and payoff depends on worker age and whether it is the current firm. This translates into three basic conditions for each period. First, the current firm's wage offer must maximize profits. Second, the rents of being the current firm are the expected present discounted

¹⁰In doing so we lose some information about the exact employment history of the worker. However, we are still able to discuss when he switches employers and his wage dynamics.

value of worker i 's product minus his wages at that firm. Third, the outside wage offer must ensure expected zero profits for the winning firm. We define these conditions formally below.

Definition 3 *The equilibrium of this economy is a series of current firm wages $\{w_{i,t}\}$, outside wages $\{\bar{w}_{i,t}\}$, and expected profits $\{V_{i,t}\}$ such that $\forall t$*

1. $w_{i,t} \in \arg \max_w (a_{i,t} + \delta V_{i,t+1} - w) [1 - G(\bar{w}_{i,t} - w - \sigma)]$ (*Profit Maximization*)
2. $V_{i,t} = \max_w (a_{i,t} + \delta V_{i,t+1} - w) [1 - G(\bar{w}_{i,t} - w - \sigma)]$ (*Expected Profits*)
3. $\bar{w}_{i,t} = a_{i,t} + \delta V_{i,t+1}$ (*Zero Profit for Outside Firms*)

We now solve for the equilibrium of this game. Substituting the Zero Profit condition into the Profit Maximization condition we get,

$$V_{i,t} = \max_w (a_{i,t} + \delta V_{i,t+1} - w) [1 - G(\bar{w}_{i,t} - w - \sigma)] \quad (2.5)$$

$$= \max_{x_{i,t}} (x_{i,t}) [1 - G(x_{i,t} - \sigma)] \quad \text{where } x_{i,t} = a_{i,t} + \delta V_{i,t+1} - w. \quad (2.6)$$

x is the ex post rent that a firm receives from the worker staying for another period. We see that the optimal value of x is tied directly to the distribution of λ .

The first order condition of (2.6) in terms of $x_{i,t}$, are,

$$x_{i,t} = \frac{1 - G(x_{i,t} - \sigma)}{g(x_{i,t} - \sigma)} = \frac{1}{H(x_{i,t} - \sigma)}. \quad (2.7)$$

Recall that the hazard rate $\frac{g}{1-G}$ is non-decreasing. The objective function is quasi-concave so the first order condition is necessary,¹¹ sufficient and has a unique

¹¹ $x_{i,t} \leq 0$ cannot solve the maximization problem since positive profits are possible. Therefore the solution to the maximization problem must be on the open set $(0, \infty)$. If the maximum exists it must satisfy the first order condition.

solution. Let x^* solve the first order condition (2.7). Since the distribution of the λ 's are stationary, the equilibrium satisfies $x_{i,t} = x^*$ every period. From (2.5) and (2.6) we know that $V_{i,t}$ must also be stationary.

$$V_{i,t} = \max_w (a_t + \delta V_{i,t+1} - w) [1 - G(\bar{w}_{i,t} - w - \sigma)] \quad (2.8)$$

$$= x^* [1 - G(x^* - \sigma)] = V^* \text{ for all } t. \quad (2.9)$$

Plugging the stationary V 's and x 's into our equilibrium definition 3 we have the unique equilibrium.

Proposition 11 *The unique equilibrium of the above economy is*

$$w_{i,t} = a_{i,t} + \delta V^* - x^* \quad (2.10)$$

$$\bar{w}_{i,t} = a_{i,t} + \delta V^* \quad (2.11)$$

$$V_{i,t} = V^* \quad (2.12)$$

$$\text{turnover probability} = G(x^* - \sigma). \quad \forall t \quad (2.13)$$

$$\text{The economic cost of employee turnover} = x^* [1 - \delta (1 - G(x^* - \sigma))] \quad (2.14)$$

Proof. above ■

Proposition 11 gives us the turnover probabilities and wage offers for worker i 's lifetime. The stationarity of this result comes from the stationary nature of the taste shock λ . The probability of turnover each period is always $G(x^* - \sigma)$. Worker i 's wages are his marginal product shifted by one of two constants depending on whether he switched employers that period. Now that we have solved this preliminary model we discuss some important aspects of the solution.

2.3.2 Wage gains at voluntary turnover

Once the worker has been employed for one period, the current firm offers a wage below marginal product.

$$w_{i,t} = a_{i,t} + \delta V^* - x^* = a_{i,t} + x^* [\delta (1 - G(x^* - \sigma)) - 1] < a_{i,t}. \quad (2.15)$$

Since the worker stays with positive probability, this creates a positive expected profit for the current firm. When hired by a new firm, the discounted value of these expected future profits are part of a worker's wage. This causes a wage premium when the worker switches firms. The wage premium in this model can also be interpreted as signing bonuses, moving reimbursement etc.

It has long been observed that voluntary turnover is associated with wage gains (ex DePasquale and Lange 1971, Topel and Ward 1992). This is not the first theoretical paper to generate this result. For example, Greenwald (1986) shows that expected future rents from asymmetric information can increase wages at turnover above expected marginal product.¹²

2.3.3 Cost of Turnover

A worker's knowledge about his feelings toward his current firm creates rents for that firm. When worker i leaves his current firm, the firm loses all claims to future rents from that worker. Proposition 1 gives us the economic loss of the firm relative to worker i staying with the firm. We now discuss why firms will be willing to incur such a loss without increasing wages up to the worker's marginal product.

¹²In his model new employers anticipate future information rent as ability is learned by only the new firm. The competitive market causes such rents to be appropriated to the worker at the beginning of the relationship.

The decision to quit for the worker is driven by relative wage offers as well as the realization of the taste shock. Since the worker's tastes are private information the current firm doesn't know exactly how much rent it can extract from the worker. The firm balances the turnover effects of offering a low wage with the profit of retaining an employee at a low wage. If the current firm were to increase its wage offer above the equilibrium wage the worker would be more likely to stay. However by so doing the firm would reduce his expected profits (since equilibrium wage offers were chosen optimally). In essence, knowledge of the worker about the current firm creates rents for the current firm. Asymmetric information between worker and firm induces turnover and wage offers that lie below the marginal product.¹³

2.3.4 Efficient turnover

The efficient turnover rule for this model is that the worker should switch employers whenever the taste parameter is less than minus the switching cost ($\lambda_{i,t}^{k_{i,t}} < -\sigma$). An economy where the current firm and the worker know the worker's taste parameter would result in this rule.¹⁴ However under asymmetric information the worker will switch firms whenever $\lambda_{i,t}^{k_{i,t}} < x^* - \sigma$. The worker switches firms too frequently from a social welfare perspective. We have discussed how asymmetric info results in turnover costs for the firm. But asymmetric information also induces a welfare loss by creating excessive turnover compared to full information.

¹³One might argue that such turnover costs can be eliminated with counteroffers. Although not included counteroffers in this case would change nothing in this model. Since the worker gets paid his marginal product by the outside market the wage decision by the current firm is the same regardless of when or how many times it is offered.

¹⁴Any positive match specific rents ($\lambda_{i,t}^{k_{i,t}} + \sigma$) would be shared somehow between the worker and the firm.

2.3.5 Turnover rate dynamics

This simple model fails to generate two well established stylized facts relating to turnover (see Mincer and Jovanovic 1981, Topel and Ward 1992). Here we mention them briefly and discuss how the model might be modified to produce these results. The first stylized fact is that turnover decreases with age. This is easily generated by assuming that switching costs increase over the worker's lifetime. This assumption might be reasonable as worker's get older they might feel a need to establish 'roots'. It might be increasingly difficult to move family away from friends or schools etc. Another possible alteration to the model is to make it a finite period framework. As workers get closer to the end of their working life the incentives to switch firms will decrease. That is, end of life effects will reduce the turnover rate over time.

The second stylized fact is that turnover decreases with firm tenure. This can be generated by including some firm specific match quality or switching costs that grows over the life of the match. This can either be an increase in productivity (ex firm specific human capital) or the worker "acquiring a taste" for the firm's unique culture.

While we are able to generate these turnover dynamics by extending the model in the aforementioned ways, our focus is on the relationship between job assignment and turnover. Therefore for clarity and tractability we omit such dynamics here.

In section 2.5 we relax the assumption about the stationary nature of the taste shock. We discuss how some of these established turnover dynamics can be generated in that extension.

2.3.6 Switching costs

Switching costs are frictions that improve the current firm's relative attractiveness to its employees. Positive switching costs are not necessary for most results of this model. However positive switching costs are a natural and realistic assumption that allow for realistic levels of turnover.

Proposition 12 (*Comparative Statics of σ*) *Given the above economy the equilibrium is affected by the switching cost (σ) in the following ways.*

- 1) $\frac{dx^*}{d\sigma} = \frac{x^* H'(x^* - \sigma)}{H(x^* - \sigma) + x^* H'(x^* - \sigma)} \in [0, 1)$
- 2) *Turnover decreases with σ*
- 3) $\frac{dV^*}{d\sigma} = 1 - G(x^* - \sigma) < 1$ *and is increasing with σ*
- 4) *Outside wage offer $\bar{w}_{i,t} = a_{i,t} + \delta V^*$ increases with σ*

Proof. in appendix ■

As the switching cost σ increases, the outside market becomes less appealing to the worker. The current firm is able to capture additional rents from this advantage (V^* increases). This increased profitability comes through decreased turnover and a reduced wage offer relative to the outside option (note $x^* = \bar{w}_{i,t} - w_{i,t}$ increases). Recall that all future profits are competed away in the first period of worker i 's relationship with a firm. As the value of being the current firm increases, so will the premium paid to become the current firm, i.e. $\bar{w}_{i,t}$ increases with σ . We are unable to determine the effect of an increase of σ on current firm wage offers. Since both x^* and $\bar{w}_{i,t}$ are increasing with σ , the effect of switching costs on current firm wages ($w_{i,t} = \bar{w}_{i,t} - x^*$) will be ambiguous.

It is no surprise that switching costs improve the current firm's profitability. However the turnover mechanism we introduce has interesting effects on the magnitude of this change. The effect of an increase in switching cost on expected profits (V^*) and relative wages ($x^* = \bar{w}_{i,t} - w_{i,t}$) is smaller than the change itself.

The effect of a change in switching cost will depend on the starting switching cost levels. When a worker has a low switching cost, such shocks play a small role in his turnover decision. The ability of his current firm to capitalize on an increase in the switching cost is tempered by the worker's high turnover rate. Employers of high switching cost workers are more isolated from the outside market. For such current firms, increases in switching costs have a greater effect on profitability. Therefore the effect of a marginal change in switching cost increases with switching cost (i.e. $\frac{\partial^2 V^*}{\partial \sigma^2} < 0$).

2.4 Job assignment and training costs

In this section we place structure on the worker productivity dynamics. We also introduce job assignment in a way that is consistent with earlier theoretical work. Particularly, we assume that jobs are ordered by responsiveness to worker ability (ex. Sattinger 1975, Rosen 1982, and Waldman 1984b).

An important new aspect of this model is task specific training costs. In this paper, training cost can arise from formal training provided by the firm. Alternatively these costs may be a short period of very low productivity caused by learning-by-doing.¹⁵

¹⁵It is likely that wages are negotiated less frequently than the early stages of the learning curve. There may be a short time where the worker is much less productive at the new job. This time of really low productivity is an implicit cost of promotion.

We find that current employers delay promotion longer than outside employers to avoid paying the training cost. This delay creates a promotion wage premium and turnover rates decrease when workers are eventually promoted.

2.4.1 Model

The only differences between this and the previous model is that we now include two jobs, a training cost and provide structure on productivity dynamics. Worker i 's effective ability is the product of their innate ability $a_i \in [a_L, a_H]$ and their labor market experience $t \in \{0, 1, 2, \dots\}$. This is¹⁶,

$$\eta_{i,t} = a_i t. \quad (2.16)$$

For each firm there are two possible jobs. If worker i performs job $j \in 1, 2$ at time t , he produces

$$y_{i,j,t} = d_j + c_j \eta_{i,t}. \quad (2.17)$$

We restrict $d_1 > 0$, $d_1 > d_2$, $c_2 > c_1 > 0$, and $a_L > 0$ to ensure a number of important productivity characteristics. First, job 1 is the efficient job assignment for all workers in their first period of work $t = 0$. Second, workers produce a positive marginal product in any period. Third, marginal product is increasing over time for all workers in all periods. Fourth, any worker's marginal product increases faster at job 2 than at job 1.

We assume that there is a training cost $b > 0$ that must be paid by a firm before the worker is able to work at job 2. Once a worker has been trained at job 2 by one firm he need not be trained again by another.

¹⁶While we assume that productivity grows linearly it is not necessary for the majority of the following results. Linearity is a sufficient condition for Corollaries 3 and 4. All other results are true under a much weaker set of assumptions on productivity growth.

We assume that ability, training, and wage offers are common knowledge. Again a worker's tastes shocks are private to the worker. The timing of each period of this game is as follows.

Stage 1: The current firm offers a wage and job assignment. If a worker has been offered promotion, the worker is trained and the current firm incurs the training cost.

Stage 2: Outside firms offer wages and job assignments.

Stage 3: Worker observes taste shock for current firm and chooses among all offers.

Stage 4: Training period for outside firms (if needed).

Stage 5: Production, wage payment, and utility experienced.

The time line is similar to Bernhardt's (1995) signaling model with the addition of the training cost. Our timing assumption creates an asymmetry between current and outside firms. The worker's ability to leave after training creates a possible loss to the current firm.¹⁷ Outside firms do not face this same risk. The worker is not able to leave a new employer until the next period.

One might be concerned that this additional asymmetry might be the cause of the later results. However this result occurs as an equilibrium under other timing assumptions. For example suppose that the switching cost is weakly greater than the training cost (i.e. $\sigma \geq b$). The only way the worker would switch again between outside firms is if the wage increase is high enough to incur the switching cost again. Under perfect competition, the increase in outside offers after training

¹⁷The worker cannot commit to stay with the firm after being trained by paying a bond at the beginning of the period (see Carmichael 1985; and Baker, Jenson and Murphy 1988).

is exactly b . The worker would never have a reason to switch even if he could. In this case the results of this model are identical to the model just presented.¹⁸

2.4.2 Equilibrium

Before we discuss the equilibrium of this model we alter some notation in order to deal with the new complications. Since firms are ex ante identical, expected future utilities still do not depend on the firm. However future expected utilities and profits may differ depending on promotion status. We denote $U_{i,j,t}$ as worker i 's expected utility at time t given that he worked at job j in period $t - 1$. $V_{i,j,t}$ is the expected profit of the period t current firm for hiring worker i in job j in period $t - 1$. As before, these values are derived under the assumption that worker i doesn't yet know any period t taste shocks.

If worker i accepts an outside offer, the worker makes his employment decision before that firm trains him. Outside firms pass any training cost directly onto the worker through a lower wage. Let $\bar{w}_{i,2,t}$ be the outside market wage offer for worker i who is assigned to job 2 and trained by an outside firm in period t . $\hat{w}_{i,2,t}$ is the outside wage offer for a previously trained job 2 worker i at time t . $\bar{w}_{i,1,t}$ is worker i 's outside market wage offer at job 1. Let $w_{i,j,t}$ be the inside firm wage offer for worker i at job $j = 1, 2$ in period t . This wage offer does not depend on the promotion history. Given the current firm's job assignment decision, training costs (if incurred) do not affect the optimal wage offer.¹⁹

¹⁸An alternative assumption is that the training is both firm and task specific. That is, training would need to happen again anytime the worker switched firms. The worker would never have a reason to quit more than once each period. Under this assumption the intuition and predictions of the model are the same although the proofs are slightly different from the version in this paper.

¹⁹If the current firm offers promotion to a worker from job 1, he must pay the training cost.

Now that we have defined the components of this game we can define equilibrium. An equilibrium of this game must satisfy four sets of conditions that define rationality for workers and firms. The first three sets of conditions are analogous to the three conditions derived in the one job model. The first set of conditions guarantee maximal worker expected utility under zero profit restrictions for outside firms.²⁰ Therefore, the wage offer for worker i must imply zero expected profits for any promotion status. Also, the outside market job assignment maximizes expected worker utility given the wages associated with each job. The second set of conditions guarantee profit maximization for the current firm. The third set of conditions define the rents of being the current firm. Such rents are the expected present discounted value of worker i 's product minus his wages at that firm. The fourth set of conditions define the worker's lifetime expected utility for each period and promotion status.

Although these conditions are a natural extension of the earlier model, the additional friction of job assignments creates a number of new variables and equations. Therefore the formal equilibrium definition is quite lengthy and is relegated to the appendix.

Proposition 13 *The above economy has a unique equilibrium with the following characteristics*

The firm's optimal wage offer solves,

$$\max_w (y_{i,2,t} + \delta V_{i,2,t+1} - w) [1 - G(\hat{w}_{i,2,t} - w - \sigma)] - b. \quad (2.18)$$

If the worker has been trained previously the firm's optimal wage offer solves,

$$\max_w (y_{i,2,t} + \delta V_{i,2,t+1} - w) [1 - G(\hat{w}_{i,2,t} - w - \sigma)]. \quad (2.19)$$

Therefore, the maximization problem for the current firm given that it trains the worker this period is the same as if the worker was trained previously (shifted by a constant).

²⁰As before this is implied by the Bertrand like competition that we assume.

- 1) $V_{i,2,t} = V^*$, $w_{i,2,t} = y_{i,2,t} + \delta V^* - x^* \quad \forall t$.
- 2) $\bar{w}_{i,2,t} = y_{i,2,t} + \delta V^* - b$, $\hat{w}_{i,2,t} = y_{i,2,t} + \delta V^* \quad \forall t$.
- 3) *If the outside firms do not offer promotion at t , then the current firm will not offer promotion. $V_{i,1,t} = V^*$ and $w_{i,1,t} = y_{i,1,t} + \delta V^* - x^*$.*
- 4) *If $y_{i,2,t_o} - y_{i,1,t_o} > b$ for some t_o then outside firms offer promotion at t_o .*
- 5) *If $y_{i,2,t_I} - y_{i,1,t_I} > (1 - \delta)b + \hat{\phi}^b$ for some t_I then current firm offers promotion at t_I . Where $\hat{\phi}^b > 0$ solves,*

$$\max_x x \left[1 - G(\hat{\phi}^b + x - \sigma) \right] = V^* - b. \quad (2.20)$$

- 6) *if $y_{i,2,\bar{t}} - y_{i,1,\bar{t}} < (1 - \delta)b + \hat{\phi}^b \leq y_{i,2,\bar{t}+1} - y_{i,1,\bar{t}+1}$ for some \bar{t} and worker i has not been promoted by \bar{t} then the current firm will deny promotion at \bar{t} . Also, if $y_{i,2,\bar{t}} - y_{i,1,\bar{t}} > b$ then turnover rates at \bar{t} are strictly greater than after promotion (at $\bar{t} + 1$).*

Proof. In appendix ■

Proposition 13 states that current and outside firms offer promotion in finite time. Worker i will be promoted by the current firm or outside market each with positive probability. Once promoted, the wage offers are analogous to the one job model. Outside firms offer promotion when the productivity at job 2 is at least b (the training cost) higher than at job 1. The current firm offers promotion whenever the worker is at least $(1 - \delta)b + \hat{\phi}^b$ more productive at job 2. However both the outside and current firms may offer promotion before this time.

We now discuss why the current firm never offers promotion before the outside market. We first explore the promotion decision for the outside firms. The outside firms offer wages and job assignments to maximize worker utility

under the zero profit condition. Outside firms directly pass the training cost on to the worker through a lower wage. Therefore outside firms promote the worker when the worker is more productive at job 2 by at least the training cost (*i.e.* $y_{i,2,t} - y_{i,1,t} > b$). It may actually happen before this depending on discounted future utilities and firm profits for the promoted vs. un-promoted worker.

If the outside market begins to offer promotion, promoting is in some sense more efficient than not promoting.²¹ Unlike the competitive outside market, the current firm maximizes profit rather than efficiency. Once the current firm decides to train, its profit will be the same as in the one job case minus the training cost (*i.e.* $V^* - b$). If the current firm does not offer promotion it can avoid paying the training costs and optimally responds to outside offers. If promotion is efficient and the current firm does not promote, outside offers improve relative to the worker's marginal product at the current firm. This increased outside competition reduces expected profits for the current firm. However since it can avoid training costs the current firm may still want to inefficiently delay promotion. The current firm delays promotion as long as the effect of outside competition on profits is small. Over time the outside pressures increase until eventually the current firm also offers promotion. It is only the relative increase in the outside wage from efficient promotion offers that causes the current firm to offer promotion as well. Therefore the current firm will never offer promotion before the outside market.

The parameter $\hat{\phi}^b$ defines the promotion rule for the current firm. It represents the efficiency gains from promoting required for the current firm to pay the training cost b . Naturally this is increasing with the training cost. Further it will be zero when the training cost is zero. Any increase in the training cost will translate into

²¹Precisely this means that current productivity and future utility for the promoted worker is at least the training cost greater than a similar un-promoted worker.

a larger gain in $\hat{\phi}^b$.²²

The behavior of the worker depends on more than just current wages, switching cost, and taste shock. Future utility is important and depends greatly on promotion status. If the worker believes that he will be denied promotion in the next period, he might prefer promotion this period at a lower present wage than to stay at job 1. These considerations make the last result of proposition 13 somewhat weak. The result says that the current firm will not offer promotion in the period just before worker i becomes $(1 - \delta)b + \hat{\phi}^b$ better at the higher job. At this time if he is at least b more productive at job 2 turnover rates will be high compared to after promotion. However, it may be the case that the current firm offers promotion well before this time. Off the path of play future utility considerations become important and can influence promotion timing.²³

We now see that the promotion decisions become cleaner when the discount factor is small compared to the difference in productivity growth between jobs. This assumption ensures a stronger result concerning the relationship between turnover and promotion.

Proposition 14 *Suppose δ is sufficiently small. Let $T_i \in \mathbb{R}$ solve $y_{i,2,T_i} - y_{i,1,T_i} = (1 - \delta)b + \hat{\phi}^b$. then the following is true in equilibrium.*

1) *The current firm offers promotion if and only if $t \geq T_i$.*

²²Recall from 2.20 that $\hat{\phi}^b$ solves $\max_x x [1 - G(\hat{\phi}^b + x - \sigma)] = V^* - b$. Therefore $\max_x x [1 - G(0 + x - \sigma)] = V^* - 0$ implies $\hat{\phi}^0 = 0$. We know from the implicit function theorem that $\frac{d\hat{\phi}^b}{db} = \frac{1}{1 - G(\hat{\phi}^b + x(\hat{\phi}^b) - \sigma)} > 1$. Therefore $\hat{\phi}^b > b$ whenever $b > 0$.

²³For example, suppose the worker knows that he will not be promoted by the current firm in the next period. Future utility considerations become meaningful for the worker. The difference in expected utility of a promoted vs. non-promoted worker may become significant enough to induce the current firm to start offering promotion earlier. That is, the current firm may offer promotion to worker i in some period t , but if the worker had not been promoted by period $t + 1$ he would not be offered promotion until period $t + 2$.

- 2) There exists a $\bar{T}_i \leq \frac{b-(d_2-d_1)}{a(c_2-c_1)}$ such that outside firms offer promotion if and only if $t \geq \bar{T}_i$.
- 3) For every $t' \in (\bar{T}_i, T_i)$ turnover of an un promoted worker i at time t' is greater than after promotion.
- 4) if $\bar{T}_i \leq t' < t'' < T_i$ turnover of an un promoted worker is greater at t'' than at t' .

Proposition 14 gives us the result that the current firm will only promote when the worker is $(1 - \delta)b + \hat{\phi}^b$ more productive at job 2. Similarly outside firms will begin promoting the worker every period after some \bar{T}_i but never before. The turnover rate for worker i increases during the time where promotion comes only from outside firms. Worker i 's turnover rate falls back to its lowest level after he is promoted.

The condition on the discount factor for this proposition relates to the length of time that workers are in the promotable state (i.e. when $(1 - \delta)b + \hat{\phi}^b > y_{i,2,t} - y_{i,1,t} > b$). If worker i will quickly transition out of this state or has a low discount factor, future expected utility differences ($\delta U_{i,2,t+1} - \delta U_{i,1,t+1}$) will be small. As discussed previously, future expected utility differences can speed up inside promotion before $y_{i,2,t} - y_{i,1,t} \geq (1 - \delta)b + \hat{\phi}^b$. If δ is small the effect of future utility is weak enough that the current firm will not promote before this rule is satisfied.

2.4.3 Turnover rate dynamics

Propositions 13 and 14 give us the empirically consistent result concerning promotion and turnover. For any worker, turnover will be weakly decreasing with

promotion. But suppose worker i experiences a period in his work life where he is at least b but less than $(1 - \delta)b + \hat{\phi}^b$ more productive at job 2 than job 1. Proposition 14 states that his average expected turnover rate will be less at job 2 than at job 1. Suppose now the population of workers is sufficiently rich that a positive proportion of workers experience a period where they are offered promotion from only outside employers. Then the average turnover rate for the population will decrease with promotion.

2.4.4 Wage Dynamics

The productivity jump at inside promotion translates into a large wage increase for the worker.

Corollary 3 *Suppose there exists some period where only the outside firms offer promotion. For a worker who is promoted within a firm, wage increases at promotion are higher than wage increases the period before or after promotion. Wage increases after promotion are higher than wage increases before promotion*

Inefficient job assignment prior to promotion drives Corollary 3. This inefficiency arises when only the outside market offers promotion. As the current firm delays promotion, the worker's current firm wage offer grows slower than his outside offers (and job 2 marginal product). When he is finally promoted, his productivity at the current firm grows by a substantial amount. His current firm wage offer jumps back to a level that is closer to the outside market wages. Since he was in some sense underpaid in the previous period, he experiences a large jump in wages when promoted from within.

This result is somewhat surprising when compared to a similar economy with no training costs. Without such costs, linear ability growth implies the wage increase at promotion is a linear combination of the wage growth the periods before and after promotion. In this case promotion always happens in the period when ability crosses the threshold of efficient promotion. Wages are marginal product minus a constant. The increase in productivity in the period of promotion is the increase in ability up to the threshold (θa_i for some $\theta \in [0, 1]$) times job 1 productivity responsiveness to ability (c_1) plus the increase in ability above the threshold ($(1 - \theta)a_i$) times job 2 productivity responsiveness to ability (c_2). This wage increase will always be less than wage increases after promotion which are always the increase in ability times job 2's responsiveness to ability ($a_i c_2$).

Corollary 4 *The wage increase for an internally promoted worker increases with the difference between his possible job 1 and job 2 productivity levels in the period prior to promotion.*

The wage increase at promotion depends on the inefficiency the current firm tolerates before offering promotion. The more inefficient the job assignment was before promotion, the greater the wage gain when finally promoted.

2.4.5 Switching Cost and Promotion

In section 2.3.6 we discussed the effect of switching costs on wage offers, turnover and profitability. We now examine the role of switching costs on the internal promotion decision.

Recall that the profit of the firm is increasing and convex in the switching cost.

The switching cost is in some sense a measure of how isolated the current firm is from outside pressures. As the switching cost increases the firm becomes more isolated. Any changes in the relative attractiveness of the wage offer has a greater effect on profitability.

Corollary 5 *The threshold of promotion ($\hat{\phi}^b$) is decreasing in the worker's switching cost.*

Proof. In appendix ■

Proposition (5) is written in terms of the 'threshold of promotion' $\hat{\phi}^b$ rather than a direct statement about turnover probability. Without knowing the exact productivity dynamics it is difficult to translate this into a precise statement on promotion in the most general cases. However, under conditions satisfying Proposition (14) we easily see that decreasing $\hat{\phi}^b$ speeds up worker's promotion (i.e. $y_{i,2,t} - y_{i,1,t} \geq (1 - \delta)b + \hat{\phi}^b$ sooner). In essence, less risk is involved with training a worker with a higher switching cost. Therefore, the current firm will train a worker with a lower turnover rate sooner. From Corollary 4 we also know that the wage premium at promotion will be smaller.

There is some preliminary empirical evidence that is consistent with this prediction. Devaro and Waldman (2006) find that the probability of promotion increases with education in the Baker et. al. (1994) data. They also find that higher education lowers the wage increase at promotion. Our model suggests that workers with a lower turnover rate are promoted earlier and will have a smaller wage increase at promotion. Therefore, their empirical results are consistent with this model if (in their data set) education is negatively correlated with the turnover rate.

2.5 Extension

It may seem unrealistic that a worker's tastes are independent and identically drawn each period. In this extension we incorporate a simple version of the standard Jovanovic (1979a) style matching model into our framework. We do not provide all the details or analysis here as much of it is similar to the earlier work. But we do provide an outline and the basic results.

We retain all of the assumptions about the taste parameter λ and job assignment from the earlier sections. There is an additional taste shock that is constant for the life of a match. We denote this as $\mu_{i,t}^k$. At the beginning of a match $\mu_{i,t}^k$ is drawn from a common distribution. After the first period of work within a firm, both the worker and the firm perfectly observe $\mu_{i,t}^k$. If the worker leaves a firm and returns at a later period he draws a new μ . Each period the worker will make its turnover decision based on the wage offer and that period's match $(\lambda_{i,t}^k + \mu_{i,t}^k)$. The firm offers a wage that maximizes its expected profit given the constant part of the match μ . The outside firms offer job assignments and wages to maximize worker expected lifetime utility subject to the zero-profit condition.

As before, this game has a unique solution with the same characteristics as the earlier models. There are a few new results related to the nature of μ . The turnover rate is decreasing in μ . Also, the worker's value to the firm (V) and the worker utility (U) will be increasing with μ .

Adding a persistent component to the worker taste shock generates the classical empirical observation of decreasing turnover rate with firm tenure (see Mincer and Jovanovic 1981; Topel and Ward 1992). In this model for a given μ turnover rate is constant. However each period, workers with lower μ 's are more likely to quit

than workers with higher μ 's. For a population with the same firm tenure, the distribution of μ moves to the right over time. Since turnover is decreasing with μ , this means that average turnover rates will decrease with firm tenure. This logic is similar to Jovanovic (1979a) in that the good matches are more likely to persist and create lower turnover in the future.

We can also see that average turnover will be decreasing in labor market experience. This follows from the above conclusion. The longer that a worker has been in the market the more likely he has found a higher persistent match quality (μ). Over time the distribution of a worker's possible μ moves to the right. However, controlling for firm tenure this effect goes away.

Our final result is that workers with a higher μ are offered promotion from the current firm earlier. This result is similar to Proposition 5 dealing with switching costs. In both cases, workers with lower turnover rates are promoted earlier than those with higher turnover rates.

2.6 Discussion

We have seen that this stylized model produces empirical predictions consistent with studies on training costs and promotion. We now discuss a few of these predictions as well as some alternative explanations.

2.6.1 Promotion and Turnover

We have discussed how training costs and our turnover assumptions can generate a negative relationship between turnover and promotion. This is consistent with empirical work discussed previously by Carson et al. (1994) and Saporta and Fajourn (2003).

Another possible explanation for decreasing turnover rates at promotion might be the tournaments argument (Lazear and Rosen 1981). The logic is as follows. Suppose that a worker has won a promotion. This is effectively the right to a wage that is above the market wage. Once promoted, the worker will be less likely to quit his current job and lose his relatively high wage.

To our knowledge, such a model has not been formally developed. We have argued that workers with a high switching cost (or high persistent taste) will be promoted earlier. Equivalently, we predict that low turnover predicts early promotion. A tournament model with endogenous turnover may generate different predictions on the type of worker that is more likely to be promoted. Tournament models use wages above marginal product as prizes to induce greater worker effort at lower levels. All else equal, firms may want to promote workers with a high turnover rate to lower the expected cost of overcompensation. However workers with low turnover may exert more effort and will be more likely to win the prize. It is difficult to say which of these effects will dominate without formulating such a model.

An empirical test that could distinguish these theories would examine turnover rates for workers who are likely to be promoted soon. In a tournament model, likely future internal promotion reduces the worker's incentive to quit. In our

model, workers who expect to be promoted soon are already receiving promotion opportunities from the outside market. Therefore a tournament model would likely predict a decreasing turnover rate before promotion. Our model predicts the turnover rate will increase.

2.6.2 Wage premium at promotion

Empirically the large wage increase at promotion has been observed by a number of studies (ex. Gerhart and Milkovich 1989; McCue 1996; Baker et al. 1994). As noted in the introduction, this paper is not the first to explain this phenomena. However, to our knowledge this is the first to generate such a result under symmetric information about worker ability and deterministic ability growth.

The signaling models of promotion (Waldman 1984b, Bernhardt 1995) generate promotion wage differentials through signaling. A single wage is paid to workers with different ability types but the same promotion history. This result largely does not fit the data. Even though job assignment is important in determining wages there is still a wide variety of wages offered to workers with the same job history and set of observable characteristics.²⁴

Tournaments (ex. Lazear and Rosen 1981) are a classic explanation for the wage increase at promotion. Workers compete for a large pay increase through investment in human capital. Therefore when workers are promoted they are given a substantial raise. While this model captures some wage and promotion dynamics there are other labor market observations not addressed by tournaments

²⁴But possibly this could be explained by outside firms having some private information on worker ability or a mix of symmetric and asymmetric learning (see Pinkston 2008, and Waldman 2008).

that are captured by other models. These include, cohort effects, wage increases predicting promotion, and wage increases upon promotion are smaller than the difference in average wage levels across the jobs.

Gibbons and Waldman (1999,2006) generate a model where wages are closely tied to ability. Wage premium at promotion are generated through their learning mechanism. However because the rate of learning slows over time, the wage premium is smaller with each successive promotion. This does aspect not fit with some empirical findings of Baker et al. (1994). They find that the wage premium at promotion is higher at later promotions.

In our model, the wage differential is tied to training costs associated with that promotion. We do not model more than one job change in this paper. However it is likely that successive job changes have higher training costs. In our model the wage premium at promotion is tied to these training costs.²⁵ Therefore, our framework can generate increasing promotion wage differentials.

2.7 Conclusion

Over the last few decades a number of theoretical models have been developed to help better understand the relationship between the worker and the firm. However little is understood about how the inner dynamics of the firm relates to worker turnover. A small set of models have emerged that model turnover by including a private worker taste (or disutility) shock. This assumption creates interesting new insights and provides non degenerate turnover in equilibrium. To our knowledge,

²⁵Again these might be higher explicit costs of training or the implicit costs of learning by doing at a more important job.

little is known about this mechanism itself. As this type of modeling is likely to become far more prevalent in the future, it is useful to better understand this turnover mechanism.

In this paper we begin to explore the implications of including such assumptions into personnel economics models. We develop a theoretically rich infinite horizon model. This model provides a theoretical foundation for the 'cost of employee turnover'. We then discuss the various forces that influence current wages and job assignment. Aspects of the labor market that have not been important now become very significant. In particular, task specific training related to job assignment can create frictions affecting turnover rates and wage and promotion dynamics

The turnover model in this paper generates a number of results consistent with empirical stylized facts. Further theoretical and empirical work will determine if this new style of modeling improves understanding of the worker and firm relationship.

CHAPTER 3

EMPLOYEE BONDING AND TURNOVER EFFICIENCY

3.1 Introduction

Various types of employee bonding are common practices in business. Institutions often provide an up-front payment or service with a written agreement that the workers remain working at the institutions for a specified time. These payments and services are as varied as signing bonuses, general skills training (i.e. college or graduate school) or moving expenses. Other types of bonding include pensions where workers only receive payment if they remain at a job for a sufficiently long period of time. While these types of contracts are common, little attention has been given to their effect on turnover efficiency. In this paper I discuss how bonding contracts can improve welfare through a reduction in excess employee turnover. I show that such contracts are only effectual if payments upon quitting are returned to the central organization or other third party and not the manager who sets the employee's wage.

In this paper I develop a two period labor turnover model to examine the usefulness of employee bonding. I assume perfect competition between firms for workers. Any expected future profits are given to the worker at the beginning of the relationship. Therefore any policy that improves efficiency improves worker utility and is preferred by the workers.

The source of turnover and efficiency in our model is (firm and time specific) private random utility shocks experienced by the worker.¹ These shocks are ob-

¹These shocks have a natural interpretation. They are the various non-pay aspects of a job whose importance to the worker is unknown to the firm. We assume that these shocks change

served only by the worker. Also these shocks are only known about the worker's current employer. The current employer is able to exploit these shocks through wage offers that are below the employee's marginal product. Therefore, the employer captures rent from the worker as long as the worker 'enjoys' working at the firm enough to compensate for any differences in wage. If the worker receives a shock that is low enough, he will quit the firm and work elsewhere.

The competitive outside market sets a wage equal to the worker's marginal product. Because the worker's current employer sets the wage below marginal product, workers change employers more frequently than is efficient. Some of the time the value of the shock implies that the worker should stay, yet it does not overcome inside and outside wage differences. I argue that this type of turnover inefficiency can be improved by employee bonding.

In order for the bonding to reduce inefficient turnover, the contract must have certain characteristics. In examining the turnover effects of such bonds we first consider a situation where the wage setter claims the payment if the employee quits. I show that under this situation, any such bond is entirely deducted from the wage of the employee. In equilibrium, turnover is identical to what would happen without any such bond.

I then consider an alternative contract where the bond is centralized away from the manager who sets the wage. In this case the wage setter does not claim the repayment if the employee quits. When setting the wage, the manager benefits from this bond creating an additional switching cost for the worker. This induces the manager to reduce the wage to some extent. But, since the manager does

over time. Workers can only predict future shocks for their current employer one period ahead. Since the worker has information about the current employer, the current employer is able to exert a kind of monopsonistic power over the employee.

not get the repayment he does not reduce the wage by the entire amount of the bond. Since the cost of repaying the bond is greater than the reduction in wage the worker is less likely to quit.

This type of contract can reduce turnover inefficiencies where without the bond the turnover rate is inefficiently high. Because the labor market is competitive any efficiency gains are passed on to the worker through higher wages in the first and/or second period. Therefore, firms that offer bonds that maximize efficiency will be preferred by potential employees. Such bonding not only reduces turnover, but is also preferred by employees.

I later discuss bonds attached to firm sponsored general skills training. The motivation for such bonding contracts is much the same as in the earlier models. However, examining bonds in this setting allows me to generate predictions that I can compare with the existing literature on general skills training.

The remainder of the paper is as follows. I review the relevant literature in Section 2. I set up and analyze the simple model in Section 3. Section 4 examines the effect of bonding offered by the manager and then the firm. In Section 5 I relate this finding to employer sponsored general skills training. I discuss firm specific human capital and pensions in Section 6. Section 7 is the conclusion.

3.2 Literature Review

One main criticism of employee bonding is that such bonds could simply be taken from the employee's wage (see Carmichael 1985). However there is evidence concerning both pensions (see Allen, Clark and McDermed 1993) and firm sponsored

training (see Manchester 2010) that suggests that bonding contracts have real effects on turnover. These papers will be discussed in more detail below in relation to the broader literature.

A large portion of the literature on employee bonding and non marginal product wages revolve around mitigating problems with moral hazard (see Lazear 1979, 1981, Yellen 1985, Carmichael 1985, Baker Jenson and Murphy 1987). In his classic paper on mandatory retirement, Lazear (1979) describes how wages above marginal product in the years right before retirement can induce workers to exert effort in earlier periods. If the worker were to shirk and be fired early on in their career, they lose out on higher pay later in life. The firm enforces mandatory retirement in order to achieve the efficient retirement date given this mechanism. Carmichael (1985) discusses a similar use for pension plans. The efficiency wage literature discusses how firms which set a wage above market equilibrium can induce extra effort (see Yellen 1985). If workers are fired for shirking they risk losing the higher wage job. While these papers do discuss how such policies affect firings, they do not discuss the worker's quitting decision.

Other work on bonding discusses the self selection of workers into or out of jobs with pension and other bonds (see Salop and Salop 1979). The idea is that workers who feel that they are less mobile will be more likely to take jobs with a bond. Firms that are more concerned with maintaining a stable work force will choose to employ such devices.

In their empirical work using PSID data, Allen, Clark and McDermed (1993) examine the effect of pensions on job mobility. It has long been observed that jobs covered by pensions exhibit low levels of turnover. They outline three possible explanations for this finding. The first is the sorting story (ala Salop and Salop

1979). The second explanation is that jobs covered by pensions tend to have higher compensation in general.² The third reason is the actual loss of the pension if the worker were to quit. Allen et al use PSID data to distinguish between the various reasons for this stylized fact. While they find that sorting and total compensation are important, the loss to the employee when he/she leaves has the greatest effect.

This paper discusses a potential role for formal programs in reducing inefficiently high turnover rates. With the exception of chapter 1 of this dissertation, the work on formal compensation rules within an organization has focused on improving performance by managers and employees. For example the theories on tournaments (ex. Lazear and Rosen 1981) and up-or-out contracts (ex. Kahn and Huberman 1988) rely on the firm's ability to commit to reward workers that exert high effort or choose high investments in human capital. Other theoretical work focuses on limiting the discretion of managers thus reducing employee influence activities (see Milgrom 1988, Milgrom and Roberts 1988a, 1988b) or favoritism (see Prendergast and Topel 1996).

In chapter 1, I show that formal pay rules can reduce excess turnover inefficiencies caused by asymmetric information concerning worker ability and wages as signals. I show that (as is also the case in this paper) there exists some excess turnover inefficiencies under full information about worker ability. I then show that asymmetric information about worker ability exacerbates this problem. Formal pay systems can reduce the higher turnover but not to the efficient level. In the current paper I show how bonding when there is full information on worker ability can bring turnover to the efficient level. The current paper should be seen as a compliment to this earlier work on asymmetric information on worker abili-

²In another empirical paper using data from the Survey of Income and Program Participation, Gustman and Steinmeier (1993) also find that total compensation is higher at jobs offering pensions. This is also controlling for worker characteristics.

ties. That is, achieving efficient turnover in a market with asymmetric information about worker abilities may require both bonding and formal pay systems.

In this paper, turnover plays a crucial role. There is a vast literature on various types of turnover mechanisms. The papers on job search and matching provide valuable insights into the turnover process (ex. Jovanovic 1979a, 1979b; Burdett 1978). In these classic models, turnover is a result of economic agents attempting to improve expected match quality. Imperfect information and learning gradually change worker and firm matches over time. While these models are able to generate a number of empirical predictions, they do not address the idea of employee turnover costs.³ Many such theories assume workers are paid their marginal product (ex. Jovanovic 1979a, 1979b). In this case there is no real loss to the firm when the worker leaves. In other turnover models workers are given a fixed proportion of their match specific rent (ex. Mortensen 1978). In this case firms do incur an economic loss when the worker switches employers. However a fixed sharing rule of the match specific rent may not be realistic. The ‘cost of employee turnover’ is not adequately addressed by these classic turnover models.⁴

The main difference between this paper and the classic turnover papers lies in the information assumption about match quality. In the current paper I assume that the quality of the match is known and experienced by the worker. In chapter 2 of this dissertation I use a similar assumption as I consider promotion timing,

³For example, in their study of four different hotels Hinkin and Tracy (2000) estimate the turnover cost of a front-desk associate to be from 5 to 12 thousand dollars. Over half of this figure is non-explicit “productivity” costs. See also Cascio (1999) and Wasmuth and Davis (1983) for additional examples of costs of employee turnover.

⁴Under traditional theories, when counteroffers are possible, the firm will never allow a worker to leave before offering him a wage equal to his marginal product. If workers are paid their marginal product, turnover of a single worker will not affect the firm’s profitability. Under some restrictive assumptions, existing theory can generate endogenous turnover while firms pay a wage below marginal product. For example, a model with random outside wage offers that assumes away counteroffers may have this feature (see, Munasinghe and Flaherty 2005).

wages, and turnover. In both chapters, the taste shock gives the current employer some monopsonistic power when choosing the wage offer. However, when workers decide to quit this induces an economic loss to the firm. I argue that this economic loss is the foundation for employee turnover costs. In chapter 2 I show that this dynamic monopsony power over current employees induces firms to create and maintain relationships with individual workers over time. I use this model to consider the promotion timing decision when promotion is costly to the firm and the firm experiences outside pressures. Aside from my work, a few papers use a similar turnover mechanism based on private worker utility (ex. Novos 1994, Acemoglu and Pischke 1998, Schonberg 2007, Ghosh 2007).

In this paper I discuss how employer sponsored training may be an avenue in which the firm can bond the worker. Traditionally, general skills training has been of great interest to labor economists. The early treatment of Becker (1962) argues that workers will always reap the rewards from general training. As a result, the worker must ultimately pay for such training either directly or through a lowered wage during training.

Later work on information asymmetry challenged this result. For example Katz and Ziderman (1990) discuss general training that is only observed by current employer. In this case the employer is able to capture information rents through uncertainty about worker productivity (see also Chang and Wang 1996). This logic is similar to Greenwald's (1986) discussion of adverse selection in the labor market. Acemoglu and Pischke (1998) show that asymmetric information on training is not needed in order to generate such rents. They show that information rents can induce employer training when there is full information on training but asymmetric information on ability. In a later paper they discuss other reasons why firms may

want to offer training (see Acemoglu and Pischke 1999). In general these labor market frictions "compress" the wage profile relative to actual productivity. Such wage compressions include labor unions and complementarities with firm specific skills.

As discussed by Becker (1962) the threat of a worker leaving the firm after training is the main limitation to firm sponsored training. Therefore the relationship between training and turnover is important. There have been theoretical papers that have discussed differences between US and Japanese employee training behaviors. Owan (2004) and Morita (2001) develop models with some equilibria characterized by low turnover and high firm sponsored training. Other equilibria have higher turnover rates and little firm sponsored training.

There has been empirical evidence that suggests that firms do pay for training and pass on benefits to employees. Loewenstein and Spletzer (1998) find positive wage returns to firm sponsored training. Further, they find that training provided at previous firms impact wages more than training provided by the current firm. In regards to employee turnover, Colleen Manchester (2008) finds that employee sponsored general skills training reduces turnover rates.

In this paper, it is not training per se that improves efficiency. But, paying for training may be a useful tool in which bonding can improve welfare. In a recent paper, Manchester (2010) shows that the bonding aspect of training has an additional effect on reducing turnover intention. In her paper she uses longitudinal data from MBA students from the University of Minnesota. She finds that tuition reimbursement programs reduce employees' intention to quit their current job.⁵

⁵The measure of turnover she uses is the response to the question, "What is the chance that you will voluntary quit your job in the next 12 months." She finds that this turnover intention was a positive and significant predictor of quits.

Additionally, forty five percent of such programs use "continued service requirements" after reimbursement. Under these requirements the worker must remain employed with the firm for a specified time or they must repay the bond. Twenty-two percent of workers receiving the tuition reimbursement are required to stay longer than 12 months. She finds that such bonding reduces turnover intention more than reimbursement without such bonding. It is the ability of the bonding contract to reduce turnover that motivates this paper on turnover efficiency.

3.3 The Model

In this section I introduce the turnover mechanism that drives this model. I develop intuition on the source of turnover costs and the rents associated with attracting and retaining individual workers. This initial analysis provides a benchmark for later analysis.

3.3.1 No Employee Bonding

There is free entry into production. All firms are identical. The only input is labor and it is inelastically supplied each period. A worker's career lasts 2 periods. All firms and workers are perfectly patient. All workers have marginal products that are i.i.d. from a commonly known distribution F with support $[a_L, a_H]$. I will denote worker i 's marginal product each period as $a_i \in [a_L, a_H]$. In the first period all information about ability is unknown. After a period of work the worker's ability becomes common knowledge.

Worker i 's utility consists of wages, a taste shock and a switching cost $\sigma \geq$

0 (if incurred). The switching cost is incurred only when the worker switches employers. Worker i 's taste shock for firm k at time t is denoted as $\lambda_{i,t}^k$ and is private information for the worker. All λ 's are drawn independently from a commonly known continuously differentiable distribution G . The distribution G has a zero mean and a continuously differentiable density g . The hazard rate $H(\cdot) = \frac{g(\cdot)}{1-G(\cdot)}$ is non-decreasing on its support. To ensure that turnover is sometimes efficient, I restrict the lower bound of the support of G to be strictly less than $-\sigma$. The worker only knows his second period taste shock for the firm he worked at in period 1. All other taste shocks are unknown. Since all firms are ex ante identical in the first period, the worker chooses the highest wage for his initial job. Perfect competition between potential employers ensures that second period expected profits are part of the worker's period 1 wage.

At the beginning of period 2 the current and outside firms offer wages to the worker. Given the current firm's wage offer the worker's utility if he stays with the firm in period 2 is,

$$u_2 = w_i + \lambda_i. \quad (3.1)$$

The worker incurs the same switching costs for all outside firms. If worker i enters the outside market, he accepts the highest outside firm wage offer, \bar{w}_i . In this case his expected utility in period 2 is,

$$u_2 = \bar{w}_i - \sigma. \quad (3.2)$$

I assume that, if indifferent, the worker remains with the current firm. Therefore, the worker will stay with his current employer if $w_i + \lambda_i \geq \bar{w}_i - \sigma$. The worker will leave the current firm with probability,

$$\Pr(w_i + \lambda_i \geq \bar{w}_i - \sigma) = 1 - G(\bar{w}_i - w_i - \sigma). \quad (3.3)$$

The payoff to the firm is simply the marginal product of its employees minus the wage costs. A firm's wage decision for a given worker i , depends on the first period employment relationship.⁶

The definition of equilibrium that I use is Perfect Bayesian Equilibrium. In the first period all firms are identical so the worker chooses the firm that offers the highest wage W . In the second period all firms learn the ability of the worker, but the value of λ is known only by the worker himself. All firms make wage offers and the worker chooses between offers. However the worker chooses his second period employer after observing his shock λ and all wage offers. The timing of the second period game is,⁷

Stage 1: All firms observe ability and the current firm offers a wage.

Stage 2: Outside firms observe the current firm wage offer and offer wages.

Stage 3: Each employee observes the taste shock for the current firm and chooses between offers.

Stage 4: Production, payment and utility are experienced.

In equilibrium the outside market will always bid wages up until the outside firms are making zero expected profits.⁸ The current firm knows the worker's

⁶We have not yet explicitly defined the outside option for the worker. It seems natural that wages should be greater than zero. In order to avoid uninteresting corner solutions or negative wages we impose the restriction that $a_i > x^*$ (where x^* uniquely solves $x^* \frac{g(x^* - \sigma)}{1 - G(x^* - \sigma)} = 1$). However, if we instead allow negative wages and assume that there is no outside option the results of the model are unchanged.

⁷In the benchmark model the timing is less important. Since workers enter the market with positive probability and ability is common knowledge, they will be offered their marginal product by the outside market regardless of the current firm's wage offer. However to be fully consistent with the subsequent model we specify the timing as below for both models.

⁸We know this because the current firm will never offer a wage greater than marginal product. The lower bound of the distribution of the taste shock is less than minus the switching cost. Therefore there is a strictly positive probability that the worker will enter the outside market.

outside wage offer and chooses a wage to maximize profits. This gives us the following definition of equilibrium for the second period game.

Definition 4 *The second period equilibrium of this economy for worker i is a current firm wage offer w_i , outside wage offer \bar{w}_i such that,*

1. $w_i \in \arg \max_w (a_i - w) [1 - G(\bar{w}_i - w - \sigma)]$ (*Profit Maximization*)
2. $\bar{w}_i = a_i$ (*Zero Profit for Outside Firms*)

Since the worker has information about his taste for the current firm, the current firm has some monopsonistic power. It will offer a wage below marginal product and if the worker's taste shock is high enough, the worker will accept the offer. The expected profits for the current firm in the second period is,

$$V_i = \max_w (a_i - w) [1 - G(\bar{w}_i - w - \sigma)]. \quad (3.4)$$

This gives the firm positive expected profits in the second period. I now solve for the equilibrium of this game. Substituting the Zero Profit condition into the Profit Maximization condition I get,

$$V_i = \max_w (a_i - w) [1 - G(a_i - w - \sigma)] \quad (3.5)$$

$$= \max_{x_i} x_i [1 - G(x_i - \sigma)] \quad \text{where } x_{i,t} = a_i - w. \quad (3.6)$$

x is the ex post rent that a firm receives if the worker stays for the second period. However, increasing x increases turnover. Note that the optimal value of x is tied directly to the distribution of λ .

Since the worker enters the market with a strictly positive probability, Bertrand like competition forces the wage offer to equal marginal product.

The first order condition of (3.6) in terms of x_i is,

$$x_i = \frac{1 - G(x_i - \sigma)}{g(x_i - \sigma)} = \frac{1}{H(x_i - \sigma)}. \quad (3.7)$$

Recall that the hazard rate $\frac{g}{1-G}$ is non-decreasing. The objective function is quasi-concave so the first order condition is necessary,⁹ sufficient and has a unique solution. Let x^* solve the first order condition (3.7). Since λ 's are drawn i.i.d. for all ability types, the equilibrium satisfies $x_i = x^*$ for all i . From (3.5) and (3.6) I know that V_i must also be the same for all i .

$$V_i = \max_w (a_i - w) [1 - G(\bar{w}_i - w - \sigma)] \quad (3.8)$$

$$= x^* [1 - G(x^* - \sigma)] = V^* \text{ for all } i \quad (3.9)$$

This gives us the unique equilibrium for every realized ability.

Proposition 15 *The unique equilibrium of the above economy is as follows, where x^* and V^* are as defined above.*

$$w_i = a_i - x^* \quad (3.10)$$

$$\bar{w}_i = a_i \quad (3.11)$$

$$V_i = V^* \quad (3.12)$$

$$\text{Turnover probability} = G(x^* - \sigma) \quad \forall i. \quad (3.13)$$

Proof. above. ■

Proposition 15 shows us that second period profit and turnover for any employee will be the same regardless of realized ability. The current firm will offer a wage

⁹ $x_{i,t} \leq 0$ cannot solve the maximization problem since positive profits are possible. Therefore the solution to the maximization problem must be on the open set $(0, \infty)$. If the maximum exists it must satisfy the first order condition.

that is exactly x^* less than marginal product. This uniform result comes from the i.i.d. assumption of the worker's taste shocks. If the worker stays with the current firm in period 2 the current firm earns a profit of x^* regardless of ability type. Since the difference between current and outside wage offers is always x^* , the turnover probability is always $G(x^* - \sigma)$. The expected second period profit from employing worker i in the first period is $x^* [1 - G(x^* - \sigma)] = V^*$ regardless of the realization of ability.

From this simple model I have established a few insights into the relationship between the worker and the firm. The information on how the worker feels toward their current employer is valuable to that firm. Wages optimally set below marginal product means that firms experience an economic loss when workers quit.

I now discuss how the information asymmetry between the worker and firm causes turnover to be too high in equilibrium. If the current firm could observe the worker's taste shock, the firm would extract all rents from the shock if it is efficient for the worker to remain with the firm. If the shock is greater than the negative switching cost the inside firm would pay just enough to induce the worker to stay. If the shock is too low the firm will never offer a high enough wage and the worker will quit. However in this model the firm does not know the taste shock. It must offer a wage depending only on the worker's ability. It cannot extract all rents but gains rents in expectation by setting this wage below marginal product ($w_i = a_i - x^*$). The worker receives an outside wage offer of exactly marginal product. Therefore he will switch firms whenever $\lambda < x^* - \sigma$. The socially optimal turnover rule is to quit when $\lambda < -\sigma$. The current firm's exploitation of monopsonistic power results in inefficiently high turnover.

The above inefficiencies come from the differences in the worker's outside wage

offer and the current firm's wage offer. Turnover would always be at the efficient level if the wages were always the same. The worker is always paid his marginal product from the outside market. Therefore turnover would be efficient if the worker were guaranteed that he will be paid his marginal product by the current firm as well. However as previously discussed, after the worker has worked for the firm there is an incentive to pay less than marginal product. One possible solution to this problem is contracting for future compensation at the beginning of the working relationship. However, this may be impossible if the worker's ability is unknown in the first period and realized ability is not verifiable in the second period. Another possible solution is for the firm to establish a reputation. But, this would require that both productivity and wages of previous individual employees are observable to incoming workers. This does not seem reasonable in many circumstances.

3.4 Bonding

In this section I explore how bonding can address turnover efficiencies. I show how this method does not require ability to be contractible. It only requires a separation between the agent who holds the bond (firm) and the agent who sets the wage for the employee (manager).¹⁰

¹⁰This usage of the bondholder as an outside party is related to Holmstrom's (1982) work on moral hazard in teams. In his paper the centralized firm relaxes balanced budget restrictions within a group of workers. Mechanisms that induce optimal effort are only available if there is an outside residual claimant.

3.4.1 Managers holding the bond.

We first consider a contract where the manager gives the worker a monetary or non-monetary payment valued at $b > 0$ during the first period. If the worker quits the firm in the second period he must repay the bond to the manager.

Since it must be repaid after quitting, the bond b acts as an additional switching cost to the worker. Therefore, given inside wage w_i and outside wage \bar{w}_i offers, the worker chooses to quit if,

$$w_i + \lambda_i \geq \bar{w}_i - \sigma - b. \quad (3.14)$$

As before, perfect competition in the outside market causes the market wage to be the worker's marginal product. The wage setting decision of the manager is the same as before but now includes a payment to the manager if the worker quits. In the second period the manager chooses the inside wage w_i to solve,

$$V_{i,m}(b) = \max_w (a_i - w) [1 - G(\bar{w}_i - w - \sigma - b)] + bG(\bar{w}_i - w - \sigma - b). \quad (3.15)$$

The first term corresponds to the payment that the manager gets if the worker stays multiplied by the probability that he stays. The second term is the payment that he gets if the worker quits. Rearranging this expression we get

$$V_{i,m}(b) = \max_w (a_i - w - b) [1 - G(\bar{w}_i - w - \sigma - b)] + b. \quad (3.16)$$

This gives us the unique equilibrium for every realized ability.

Proposition 16 *The unique second period equilibrium of the above economy with*

bond b is as follows, where x^* and V^* are as defined above.

$$w_i = a_i - x^* - b \quad (3.17)$$

$$\bar{w}_i = a_i \quad (3.18)$$

$$V_i = V^* + b \quad (3.19)$$

$$\text{Turnover probability} = G(x^* - \sigma) \quad \forall i. \quad (3.20)$$

Proof. Above. ■

Proposition 16 shows us that very little changes with this type of bond. Any bond b given to the worker in period one is simply deducted from his pay in period two. If the worker stays at the job his pay is exactly b less than it would have been had he not gotten the bond. If he quits then he pays the manager exactly b . Turnover probabilities are exactly what they would be without the bond.

This result illustrates that as long as wages are not predetermined, it is not simply the presence of the bond that improves employee retention. The repayment to the manager if the worker quits negates the effectiveness of the bond as an additional switching cost. I now show that without the repayment to the manager, such bonds can be effective.

3.4.2 Firms holding the bond

We now consider an alternative bonding contract that utilizes a third party. Consider a payment issued by a third party that is somewhat isolated from both the manager who sets the wage and the employee. We refer to this third party as the firm. The firm gives an amount b to the worker in period one. If the worker de-

cides to quit in period 2 the worker pays the bond back to the firm. The manager does not get any of the bond repayment if the worker quits.

Since the worker does not care whom he must pay the bond to, his quitting decision remains unchanged. However the second period wage decision is altered so that the manager no longer gets paid if the worker quits. Although the bonding contract does not directly involve the manager, the bond creates additional turnover frictions that the manager can exploit.

Proposition 17 *Under a third party bond contract as described above, second period equilibrium turnover is decreasing with the bond b .*

Proof. In Appendix ■

Proposition 17 shows that it is possible to reduce turnover inefficiencies through a bonding contract. Because the market is competitive any improvements in efficiency are passed on to the worker. The worker prefers to work for firms who are able to reduce or eliminate turnover inefficiencies. I first describe the equilibrium bond size before I discuss the intuition for the bond's effectiveness.

Proposition 18 *The unique equilibrium of the entire game has the following properties.*

- 1) *All firms offer a bond equal to $b^* = \frac{1-G(-\sigma)}{g(-\sigma)}$ to all new employees in the first period.*
- 2) *Managers choose second period wages equal to $w_i = a_i - b$*
- 3) *Equilibrium turnover is efficient.*
- 4) *First period wages equal first period productivity plus expected second period profits.*

Proof. in Appendix. ■

The most important aspect of the equilibrium described in Proposition 18 is that the bond reduces second period turnover to the optimal level. The first best turnover decision is to quit if and only if $\lambda < -\sigma$. This corresponds to a turnover rate of $G(-\sigma)$. The bond size that eliminates turnover inefficiencies has the property that the manager's ex post profit, $(a_i - w)$, is exactly equal to the bond. The wage that the manager chooses to offer the worker in the second period must maximize his objectives. The wage that induces efficient turnover will be the same as the manager's optimal wage only when the bond is of the size $b^* = \frac{1-G(-\sigma)}{g(-\sigma)}$.¹¹

Proposition 17 states that this arrangement is able to reduce turnover probabilities where the previous bond contract cannot. The key to this contract lowering turnover is that the agent setting the wage does not get compensated if the worker quits. This however does not mean that the bond doesn't increase second period profits for the manager. Naturally, as the bond affects the quitting decision of the worker, the manager is able to exploit this friction by offering a lower wage. However the resulting wage reduction is smaller than the size of the bond.

I will now explain why the manager reduces wages by a smaller amount than the bond. The key to this is the characteristics of the turnover mechanism. To illustrate this I first discuss the model without the turnover mechanism. Suppose that there is no private taste shock in this model. The employee would stay with the firm as long as the inside wage is greater than the outside wage minus the switching cost and bond. The manager wants to maximize profits and therefore

¹¹The first order condition that defines the manager's optimal wage is $a_i - w_i = \frac{1-G(a_i - w_i - b - \sigma)}{g(a_i - w_i - b - \sigma)}$. Efficient turnover happens when $a_i - w = b$. Both of these conditions are satisfied when the bond is $b^* = \frac{1-G(-\sigma)}{g(-\sigma)}$.

he chooses the lowest wage so that the employee does not quit. Therefore as the bond increases the wage offered decreases by the same amount. Without the private taste shocks it doesn't matter if the manager gets the bond if the worker quits or not. There would be no turnover in equilibrium and wages are the same regardless of who is paid the bond if the worker quits.

I now examine the wage decision when there are private taste shocks. If the worker's bond increases, the manager has the option of decreasing the wage by the same amount without affecting turnover. If he does this, the ex post profits if the worker stays is larger. Therefore, the manager incurs a greater loss if the worker decides to quit. Because of this greater loss, the manager has an increased incentive to reduce the turnover probability. The manager reduces turnover by choosing to decrease the wage by less than the amount of the bond.

The relationship between wages and later compensation is consistent with empirical work by Montgomery and Shaw (1997). In their paper they use the 1983 Survey of Consumer Finances and detailed data on pensions from the Pension Provider Survey to show a negative relationship between pension size and wages. Further the reduction in wages is smaller than the increase in the value of the pension.

The key to a bond's ability to simultaneously reduce wages and turnover depends on the structure of the contract. As we have seen in Proposition 16, turnover is unchanged if the manager is paid if the worker quits. It is only the use of a third party that allows for turnover inefficiencies.

One critical aspect not formally modeled in this paper is that the firm must have the ability not to interfere in the manager's wage setting decision after the

initial contract. That is, the firm cannot choose to reward the manager with extra compensation when the employee quits. This type of commitment seems reasonable through reputation as long as the firm is long lived.

The ability of firms to implement the necessary separation between the bond holder and the wage setter may be tied to the firm's size and structure. Only firms that are able to separate the wage setter from the bond holder will be able to use this mechanism to reduce turnover. It may seem reasonable to assume that larger firms might have more commitment power and sufficient bureaucracies to utilize such bonding.

Two characteristics of effective bonding in this model include turnover reduction and a wage reduction that is smaller than the bond. I now discuss empirical evidence on the relationship between bonding, turnover and wage reduction for different firms of different sizes. One study by Even and Macpherson (1996) uses a variety of data sets to conclude that generous pensions at larger firms reduce turnover more than pensions at smaller firms. This is of course controlling for pension and worker characteristics. Montgomery and Shaw (1997) show that the wage reduction associated with an increase in pensions are smaller for larger firms. Both of these pieces of evidence suggest that the effectiveness of bonding in reducing turnover inefficiency increases with firm size.

3.5 General Skills Training

One very common practice in business is the provision of general skills. In this section I show how bonding contracts tied to the provision of general skills training can improve turnover. As before, this turnover reduction is welfare improving and

thus preferred by workers. Further this model generates predictions on turnover and income that are consistent with empirical findings.

3.5.1 Worker Paid Training

One major prediction in the seminal paper by Becker (1962) is that it doesn't matter who contractually pays for general skills training. Once trained, the worker will be offered his marginal product by outside employers. Therefore the only way that the firm would be willing to pay for the training is if the compensation was lowered by the same amount. In this case even though the firm may be paying the direct costs, these costs are passed on to the worker who essentially pays for his own training. In order to compare the above results to that of worker sponsored training I briefly discuss general skills training by the worker.

We now consider the effect of general skills training that is paid for by the worker. The timing of the game is now as follows. After the first period of work, the worker purchases training at a cost of b . In order to isolate the effect of the bond on turnover efficiency I consider training that in itself creates no efficiency gains or losses.¹² That is, training that costs the amount b will increase the marginal product of the worker by the same amount b .

For an inside wage w_i and outside wage \bar{w}_i the worker will choose to quit the firm if $w_i + \lambda_i \geq \bar{w}_i - \sigma - b$. Because the worker gets to keep his increased productivity to the outside market, the new outside wage is $\bar{w}_i = a_i + b$. The repayment to the firm is perfectly offset by the increase in productivity.

¹²Some of the limitations on the use of bonds may be on social acceptability of different types of contracts. Minimum employment requirements tied to firm sponsored general skills training seem to be prevalent and acceptable. I use this assumption on training costs and benefits to generate predictions specific to this type of bond that are independent of training efficiency.

Proposition 19 *If workers directly pay for general skills training b , the unique equilibrium of the above economy is as follows, where x^* and V^* are as defined above.*

$$w_i = a_i + b - x^* \quad (3.21)$$

$$\bar{w}_i = a_i + b \quad (3.22)$$

$$V_i = V^* \quad (3.23)$$

$$\text{Turnover probability} = G(x^* - \sigma) \quad \forall i. \quad (3.24)$$

Proof. *Analogous to proof of Proposition 1. ■*

The importance of Proposition 19 is to point out that it is not the training that reduces turnover. Turnover rates are the same as without any bonding or training. Once the worker has paid for his own training, he is treated by the firm just as he was before only paid a higher wage.

3.5.2 Firm Sponsored Training

We now consider the effect of general skills training that is paid for by the firm. Instead of the worker directly paying the training costs, the firm pays for training and the manager sets the second period wage. If the worker chooses to quit after training but before the second period, he must repay the training cost to the firm.

For an inside wage w_i and outside wage \bar{w}_i the worker will choose to quit the firm if $w_i + \lambda_i \geq \bar{w}_i - \sigma - b$. Because the worker gets to keep his increased productivity to the outside market, the new outside wage is $\bar{w}_i = a_i + b$. The repayment to the firm is perfectly offset by the increase in productivity.

As before the manager offers a wage to the worker to maximize second period profits. Since the worker now is more productive, the manager's second period wage decision is to solve,

$$\max_w (a_i + b - w) [1 - G(a_i - w - \sigma)]. \quad (3.25)$$

Proposition 20 *The training b results in an increase in wage that is smaller than b and a reduction in turnover probability.*

Proof. In Appendix. ■

Proposition 20 states that this type of bonding results in a reduction in turnover. Even though the bond must be repaid, some of the increase in productivity is passed on to the worker through higher pay. This increase in pay reduces turnover rates. As before turnover reduction improves welfare up to a point. Therefore, the firm will want to consider the appropriate amount of training in order to minimize inefficiencies.

I now discuss how this turnover reduction is accomplished through training and bonding. The worker is required to pay training costs if he quits. Therefore the profitability increase from training could improve only the inside option. The outside option of quitting remains the same as without the training. Since the bond must be repaid, the worker's outside option is simply equal to the worker's ability without the training. Under the assumption that training improves productivity by the cost of training there is no change in the outside option..

As before, the reason the contract can increase wages and retention is because it utilizes a third party (or the centralized firm). The manager who sets the wage is not reimbursed if the worker quits. Since the worker must repay the

bond if he quits, the value of the worker's outside option is the same as without training. The manager has the option of setting the wage equal to the optimal no-training wage. If he does this, turnover rates will be the same as without training. But now, because the worker is more productive than without training, the loss associated with turnover is greater. As a result, the no-training wage is no longer optimal. The manager is willing to offer a higher wage in order to increase retention. But, the manager's optimal wage increase will be less than the productivity gain associated with training.

As with other types of employee bonding, firm sponsored training has the ability to reduce turnover in equilibrium. In the absence of such bonding, the resulting second period turnover is inefficiently high. In this model we have assumed that the labor market is competitive. We have shown that even without efficiency gains in production, training with bonds can be preferred by workers. This rationale behind training is different than those of other explanations of general training (for example Katz and Ziderman 1990, and Acemoglu and Pischke 1998, 1999). Unlike other work, it is not asymmetric information about abilities or wage compression that motivates the firm to provide such training. Providing training is simply another way that the firm has the ability to provide a turnover reducing bond.

In this model I have extracted away from asymmetric information about abilities or other wage compression explanations of general skills training. There is some evidence to suggest that training can generate such frictions (see Manchester 2008, 2010). But, there is also evidence to suggest that continued employment bonds have an additional role in reducing turnover rates (see Manchester 2010). It seems that continued employment contracts connected to the firm's provision of general skills are socially acceptable. Firms can use such bonds to reduce turnover

inefficiencies and be a more competitive employer.

3.6 Firm Specific Training and Pensions

Thus far we have discussed how bonding can overcome turnover inefficiencies caused by monopsonistic behavior of firms. The source of inefficiency is rooted in the turnover mechanism of worker taste shocks. In this section I discuss why some firms and industries may have a greater need for such bonds. I consider how the loss of firm specific human capital can be an additional source of turnover inefficiencies.

It has been shown in previous theoretical work that firm specific human capital can increase retention (see Munasinghe and O’Flaherty 2005). I show that this is also true in our model even without bonding. However even with this increased retention, turnover inefficiencies may be greater in the presence of firm specific human capital. The source of additional inefficiency is the loss of positive firm specific human capital when the worker quits. When the worker develops firm specific human capital, the manager passes some of the additional productivity on to the worker thus reducing turnover. However the manager also keeps some of this increased productivity for himself. Although turnover is reduced, it is not reduced enough to overcome additional losses from firm specific human capital. Because the manager keeps some of the increased productivity, the range of taste shocks that result in inefficient turnover increases with firm specific human capital. Depending on the distribution of taste shocks this can cause higher expected turnover inefficiencies.

The setup in the model is exactly the same as in Section 3.3 with the following

change. The worker has a higher marginal product in the second period at the firm he worked for in the first period. In the second period he produces a_i if he changes employers and $a_i + K$ if he stays with the current firm (where $K > 0$). This additional amount K represents firm specific human capital accrued over firm tenure.

Let us consider the effect of the firm specific human capital on wages and turnover in the absence of bonding.

Proposition 21 *Without bonding, equilibrium turnover rates are decreasing with firm specific human capital.*

Proof. In Appendix. ■

This result is consistent with the theoretical model of Munasinghe and O’Flaherty (2005) who find that returns made from investments in firm specific human capital reduce turnover. In essence, the firm specific human capital works in a similar way to the bond or employee switching costs. That is, the firm specific human capital is an additional benefit from the match between the worker and manager. The manager is willing to share this benefit with the worker in order to reduce turnover. Although firm specific human capital reduces turnover, there may be an increase in the economic loss associated with turnover.

Lemma 1 *The range of taste shocks that result in inefficient turnover is increasing with the level of firm specific human capital.¹³*

Proof. In Appendix. ■

¹³This assumes efficient turnover is possible. That is, the probability that $\lambda < -\sigma - K$ is strictly greater than zero.

Lemma 1 comes from the increased losses to society when a worker quits a job where he has firm specific human capital. It becomes efficient to stay with the firm for lower draws of the taste shock when there is firm specific human capital. This efficient cutoff value falls at the same rate as the firm specific human capital. The actual turnover rate is also reduced by firm specific human capital but by a smaller degree. This is because the manager does not pass on all firm specific human capital gains to the worker through a higher wage. The manager instead chooses to keep some of that additional productivity as ex post profit. That is, the manager increases the wage by a smaller amount than the increase in firm specific human capital. Therefore, the actual retention taste shock cutoff value decreases by a smaller amount than the amount of firm specific human capital. The range of taste shocks that result in inefficient turnover gets larger with firm specific human capital.

Lemma 1 does not directly consider expected turnover inefficiencies. It only deals with the range of taste shocks that result in inefficient turnover. In order to show that firm specific human capital increases expected losses from turnover one must consider the distribution of taste shocks. For example, if the distribution of taste shocks is sufficiently flat in the relevant range, higher firm specific human capital will result in greater efficiency losses. Firms facing such issues may have a greater need to employ bonds such as pensions to reduce or eliminate high turnover inefficiencies.

One interesting prediction of the model deals with the lifetime compensation differences between pensioned and non pensioned jobs in a competitive equilibrium. In a competitive labor market, ex ante identical workers can have higher lifetime compensation at jobs with pensions. The reasoning is as follows. There are two

sources of utility in this model. These are money and non pecuniary taste shocks. Jobs that offer pensions exhibit lower turnover rates. Therefore, the average taste shock of workers who stay at pensioned jobs will be lower than the average taste shock for those who stay at non pensioned jobs. In a competitive labor market, the worker is indifferent between taking a pensioned or a non pensioned job. Workers expect to experience lower non pecuniary utility at the pensioned job. If workers are indifferent between pensioned and un-pensioned jobs, lifetime monetary payment must be higher at the pensioned job in order to compensate for lower non monetary compensation. This logic is consistent with empirical work that finds that pensioned jobs exhibit higher lifetime monetary compensation for the same observable worker characteristics (see Allen, Clark and McDermid 1993, and Gustman and Steinmeier 1993).

3.7 Conclusion

This paper continues the discussion of employee turnover using a relatively unexplored and intuitive turnover device. Clearly, firms are concerned with employee retention and discussing various types of compensation in relation to turnover seems natural. If workers' quitting decisions are randomly affected by forces outside the wages offered, the question of turnover efficiency seems relevant.

In this paper I have discussed how bonding contracts are related to turnover rates and turnover efficiency. This work is quite different in motivation from other explanations addressing moral hazard considerations. Unlike the moral hazard literature this paper focuses on the quitting decision rather than firing. To my knowledge this is the first paper that has addressed the effect of such bonds

on the efficiency of the quitting decision.

In this paper I have characterized how the bond contract must be written in order to be effective. I have also discussed a number of different types of bonds seen in industry. I have related this work to empirical studies on pensions and continued employment requirements tied to tuition reimbursement plans. This model generates explanations and predictions consistent with empirical findings.

APPENDIX A

ASYMMETRIC INFORMATION AND WAGE SIGNALS:

Proof. Of Claim 1. By contradiction. Suppose there exists some $a'' > a'$ where $w_i(a'') < w_i(a')$, Let us denote $G'' = G(b_i(w_i(a'')) - w_i(a'') - \sigma)$, and $G' = G(b_i(w_i(a')) - w_i(a') - \sigma)$. Sequential rationality for type a' implies

$$(a' - w_i(a')) [1 - G'] \geq (a' - w_i(a'')) [1 - G'']. \quad (\text{A.1})$$

$\Rightarrow [1 - G'] > [1 - G'']$. Sequential rationality for a'' implies, $(a'' - w_i(a'')) [1 - G''] \geq (a'' - w_i(a')) [1 - G']$. $w_i(a'') < w_i(a')$ and $[1 - G'] > [1 - G'']$ implies,

$$\begin{aligned} (a'' - w_i(a'')) [1 - G''] &\geq (a'' - w_i(a')) [1 - G'] & (\text{A.2}) \\ &> (a'' - w_i(a')) [1 - G''] \\ &> (a'' - w_i(a'')) [1 - G''] \end{aligned}$$

Which is a contradiction. Therefore the wage offers must be weakly increasing with ability. ■

Proof. of Proposition 2 First show that the wage profile is continuous, For the type a_H , the only possible solution is $w_i(a_H) = a_H - x^*$. Where $x^* = \arg \max x [1 - G(x - \sigma)]$. This solves the first order condition $[1 - G(x^* - \sigma)] - x^* g(x^* - \sigma) = 0$. Since the hazard rate is non decreasing, the expression $x [1 - G(x - \sigma)]$ is increasing for $0 < x < x^*$ and decreasing when $x > x^*$.

I denote $x(a) = a - w_i(a)$ as the ex post rent on a worker with ability i if he stays with the firm. Given $a'' > a'$ Let us denote $G'' = G(b_i(w_i(a'')) - w_i(a'') - \sigma)$, and $G' = G(b_i(w_i(a')) - w_i(a') - \sigma)$. Sequential rationality for type a' implies

$$(a' - w_i(a')) [1 - G'] \geq (a' - w_i(a'')) [1 - G'']. \quad (\text{A.3})$$

$$\text{and } (a'' - w_i(a'')) [1 - G''] \geq (a'' - w_i(a')) [1 - G'] \quad (\text{A.4})$$

$$\begin{aligned}
&\Rightarrow (a'' - a')[1 - G''] \geq x(a'')[1 - G''] - x(a')[1 - G'] & (A.5) \\
&\geq (a'' - a')[1 - G']
\end{aligned}$$

Since $[1 - G''] \geq [1 - G']$ This implies that $x(a'') \leq x(a')$. $x(a'')[1 - G''] - x(a')[1 - G'] > 0$ and $a'' \leq a_H$ implies that $x(a'') < x(a')$. By the squeeze principle the expression $x(a)[1 - G(x(a) - \sigma)]$ is continuous with respect to a . Since $x(a)$ is strictly monotonically decreasing on $a < a_H$ and $x[1 - G(x - \sigma)]$ is strictly monotonically decreasing on $x < x^*$ therefore $x(a)$ must also be continuous as well as $w_i(\cdot)$. Now I can show that the relationship between ability types and $x(a)$ must be unique. Since $x(\cdot)$ is strictly decreasing with a it is invertible. I will now solve for the derivative of $\hat{a}(\cdot) = x^{-1}(\cdot)$. Rearranging the inequality (A.5) I get

$$\frac{a'' - a'}{[1 - G']} \left(\frac{G'' - G'}{x(a'') - x(a')} \right) + 1 - \frac{x(a'')}{[1 - G']} \left(\frac{G'' - G'}{x(a'') - x(a')} \right) \quad (A.6)$$

$$\geq \frac{a'' - a'}{x(a'') - x(a')} \geq 1 - \frac{x(a'')}{[1 - G']} \left(\frac{G'' - G'}{x(a'') - x(a')} \right) \quad (A.7)$$

As $a'' - a' \rightarrow 0$ by the squeeze theorem I know that $\frac{d\hat{a}}{dx} = 1 - x \frac{g(x-\sigma)}{[1-G(x-\sigma)]}$. Since I know $x(a_H) = x^*$ this uniquely defines the function,

$$\hat{a}(x) = a_H + \int_{x^*}^x \left(1 - z \frac{g(z-\sigma)}{[1-G(z-\sigma)]} \right) dz < a_H.$$

Therefore if there is a separating equilibrium it must be, $w_i(a_i) = a_i - \hat{a}^{-1}(a_i)$ for all $a_i \in [a_L, a_H]$ with beliefs $b_i(w) = w_i^{-1}(w)$ for all $w \in [w_i(a_L), w_i(a_H)]$. The equilibrium beliefs off of the path of play are not unique but one solution is $b_i(w) = a_H$ for all $w \notin [w_i(a_L), w_i(a_H)]$. I will now show that this is indeed an equilibrium. Given a_i and the beliefs of the outside market $b_i(w)$ the firm's decision is to find a wage offer

$$w_i(a_i) \in \arg \max_w (a_i - w) [1 - G(b_i(w) - w - \sigma)]$$

Since the beliefs are continuous and monotonic for $w \in [w_i(a_L), w_i(a_H)]$ I can show that there is no incentive for the firm to deviate to another wage on the path of

play. I will examine an equivalent problem is to choose the signaled ability by a choice of signaled x . Therefore the problem now becomes

$$\hat{a}^{-1}(a_i) \in \arg \max_{x \in [\hat{a}^{-1}(a_H), \hat{a}^{-1}(a_L)]} (a_i - \hat{a}(x) + x) [1 - G(x - \sigma)] \quad (\text{A.8})$$

The first derivative of the objective function with respect to x is

$$(1 - \hat{a}'(x)) [1 - G(x - \sigma)] - (a_i - \hat{a}(x) + x) g(x - \sigma). \quad (\text{A.9})$$

Substituting in $\hat{a}'(x)$ this becomes

$$\begin{aligned} & \left(1 - \left(1 - x \frac{g(x - \sigma)}{[1 - G(x - \sigma)]} \right) \right) [1 - G(x - \sigma)] - (a_i - \hat{a}(x) + x) g(x - \sigma) \\ &= (\hat{a}(x) - a_i) g(x - \sigma). \end{aligned} \quad (\text{A.10})$$

Since $g(x - \sigma) > 0$ always this objective function is increasing for $x < \hat{a}^{-1}(a_i)$, increasing for $x > \hat{a}^{-1}(a_i)$ and is maximized exactly at $x = \hat{a}^{-1}(a_i)$. To show that the firm cannot earn more expected profits by offering a wage off the path of play I need to show that paying a wage higher than $w_i(a_H)$ is worse than paying $w_i(a_H)$. Consider any wage higher than $w_i(a_H)$ denoted \bar{w} .

$$\begin{aligned} (a' - \bar{w}) [1 - \bar{G}] &= (a' - a_H + a_H - \bar{w}) [1 - \bar{G}] \\ &= (a' - a_H) [1 - \bar{G}] + (a_H - \bar{w}) [1 - \bar{G}] \\ &< (a' - a_H) [1 - G^H] + (a_H - w_i(a_H)) [1 - G^H] \end{aligned} \quad (\text{A.11})$$

Since $w_i(a_H)$ was chosen to maximize the profits of the a_H employer and $\bar{w} > w_i(a_H) \Rightarrow [1 - G^H] > [1 - \bar{G}]$. A Similar argument shows that the firm will never offer a wage lower than $w_i(a_L)$. Therefore this is an equilibrium. ■

Proof. of Proposition 5. Given a single wage \bar{w} , the probability of turnover is $G(a - \bar{w} - \sigma)$ which is increasing with ability (a). Since the turnover rate is increasing with ability this means that the turnover rate on a workers below average

ability is less than the turnover rate on workers above the average ability. Therefore $\int_{a_i}^{\bar{a}} (a - \bar{a})G(a - \bar{w})dF(a) - \int_{\bar{a}}^{a_H} (a - \bar{a})G(a - \bar{w})dF(a) > 0 \Rightarrow \frac{\int_{a_L}^{a_H} aG(a - \bar{w})dF(a)}{\int_{a_L}^{a_H} G(a - \bar{w})dF(a)} > \bar{a}$. The average ability of the quitter is higher than the average ability overall (and by extension the stayer) ■

Proof. of Proposition 6. Given an upper wage bound $\hat{w}_h < a_H - x^*$ induces some pooling at the top of the ability distribution. That is there exists some a'' such that all ability types greater than a'' are offered \hat{w}_h . The average ability of the worker that is offered \hat{w}_h is \bar{a}_h . Further since $\hat{w}_h = \bar{a}_h - x^*$ and $\bar{a}_h = E[a|a > a'']$, a'' and \hat{w}_h are jointly determined. By the continuity this implies that the equilibrium wage profile $w(a)$ must satisfy,

$$(a'' - w_i(a'')) [1 - G(a'' - w_i(a'') - \sigma)] = (a'' - \hat{w}_h) [1 - G(\bar{a}_h - \hat{w}_h - \sigma)] \quad (\text{A.12})$$

Defining as before $x(a) = a - w(a)$, this equation can be written as

$$x(a'') [1 - G(x_i(a'') - \sigma)] = (a'' - \bar{a}_h + x^*) [1 - G(x^* - \sigma)] \quad (\text{A.13})$$

The separation that happens below the highest wage is defined from the ability associated with each level of x . The unique unrestricted equilibrium is defined as defined as,

$$\hat{a}(x) = a'' + \int_{x_i(a'')}^x \left(1 - z \frac{g(z - \sigma)}{[1 - G(z - \sigma)]} \right) dz \quad \text{for } x > x_i(a''). \quad (\text{A.14})$$

The unique fully separating equilibria is $w_i(a_i) = a_i - \hat{a}^{-1}(a_i)$ for all $a_i \in [a_L, a_H]$ with beliefs $b_i(w) = w_i^{-1}(w)$ for all $w \in [w_i(a_L), w_i(a_H)]$. The equilibrium beliefs off of the path of play are not unique but one solution is $b_i(w) = a_H$ for all $w \notin [w_i(a_L), w_i(a_H)]$. I know that for the separated ability an increase in x decreases profitability. Naturally pooling at the top increases average profitability of highest paid workers to its maximal level. In order to show that some restriction is profit

improving it is sufficient to show that such pooling lowers the ability associated with any given x . Taking the derivative of A.14 with respect to a'' .

$$\frac{d\hat{a}(x)}{da''} = 1 - x'(a'') \left(1 - x_i(a'') \frac{g(x_i(a'') - \sigma)}{[1 - G(x_i(a'') - \sigma)]} \right). \quad (\text{A.15})$$

Taking the derivative with respect to A.13 with respect to a'' gives us,

$$x'(a'') \left(1 - x_i(a'') \frac{g(x_i(a'') - \sigma)}{1 - G(x_i(a'') - \sigma)} \right) = \left(1 - \frac{d\bar{a}}{da''} \right) \frac{1 - G(x^* - \sigma)}{1 - G(x_i(a'') - \sigma)}. \quad (\text{A.16})$$

Combining these equations I get

$$\frac{d\hat{a}(x)}{da''} = 1 - \left(1 - \frac{d\bar{a}}{da''} \right) \frac{1 - G(x^* - \sigma)}{1 - G(x_i(a'') - \sigma)}. \quad (\text{A.17})$$

Since $\lim_{a'' \rightarrow a_H} \frac{d\bar{a}}{da''} = \frac{1}{2}$ reducing a'' improves profits on nonpooled ability types.

■

Proof. Of Proposition 7. I first show that the expected welfare function given ability type a is convex with respect to a

$$\Lambda(a) = x(a) [1 - G(x(a) - \sigma)] + \int_{x(a)}^{\infty} [1 - G(\lambda - \sigma)] d\lambda \quad (\text{A.18})$$

$$\Lambda'(a) = [1 - G(x(a) - \sigma)] \left[1 - x(a) \frac{g(x(a) - \sigma)}{[1 - G(x(a) - \sigma)]} - 1 \right] x'(a) \quad (\text{A.19})$$

$$= [1 - G(x(a) - \sigma)] (1 - x'(a)) \quad (\text{A.20})$$

$$\begin{aligned} \Lambda''(a) &= -g(x(a) - \sigma) x'(a) + g(x(a) - \sigma) (x'(a))^2 \\ &\quad - [1 - G(x(a) - \sigma)] x''(a) \end{aligned} \quad (\text{A.21})$$

$$> 0$$

Because this is convex the expected welfare on workers within the range is greater than the expected welfare on the average worker within that range. $E[\Lambda(a) | a \in [a_L, a']] \geq \Lambda([E(a) | a \in [a_L, a']])$. However these two get closer if the distribution is less spread out.

Take any $a' > a_L$ and I will now show that a lower wage restriction w_L that induces pooling from a_L to a' has a lower \hat{x}_l than the $x(a)$ for the mean ability worker within that range \hat{a} .

Let $\hat{x} = \hat{a} - \hat{w}_l$ be the average expected ex post benefit to the firm. The individual rationality condition for the worker type a' is,

$$(a' - \hat{a} + \hat{x}) [1 - G(\hat{x} - \sigma)] = x(a') [1 - G(x(a') - \sigma)]. \quad (\text{A.22})$$

I now show that $\hat{x} < x(a)$ by contradiction. Suppose that $\hat{x} \geq x(\hat{a}) \Rightarrow$ the expected profit and retention rate on the average worker is less under pooling than under full separation. *i.e.*

$$(\hat{x}) [1 - G(\hat{x} - \sigma)] \leq x(\hat{a}) [1 - G(x(\hat{a}) - \sigma)]. \quad (\text{A.23})$$

Also since the profit function $V(a) = x(a) [1 - G(x(a) - \sigma)]$ is convex and increasing in a with a slope equal to $[1 - G(x(a) - \sigma)]$, The slope between the point $(\hat{a}, V(\hat{a}))$ and $(a', V(a'))$ which is must be greater than the slope at \hat{a} . Therefore $\frac{V(a') - V(\hat{a})}{a' - \hat{a}} > [1 - G(x(\hat{a}) - \sigma)]$. Rearranging this we get

$$(a' - \hat{a} + x(\hat{a})) [1 - G(x(\hat{a}) - \sigma)] < V(a'). \quad (\text{A.24})$$

Combining this with A.23 we get

$$(a' - \hat{a} + \hat{x}) [1 - G(\hat{x} - \sigma)] \leq (a' - \hat{a} + x(\hat{a})) [1 - G(x(\hat{a}) - \sigma)] < V(a').$$

This is a contradiction therefore $\hat{x} < x(\hat{a})$. Since the turnover rate on the pooled wage is now lower than the turnover rate of the average worker (within that range) under separation, if the distribution of ability types within this range is not too spread out, then the lower bound can improve welfare. ■

Proof. Of Proposition 10. Since the budget constraint is determined prior to the learning of abilities it may not be optimal for a given set of workers. I can see

this as I examine the first order conditions for the maximization problem. The Lagrangian for this problem is,

$$L(w_1, \dots, w_n, \mu) = \sum_{i=1}^n a_i [1 - G(a_i - w_i - \sigma)] + \mu (K - \sum_{i=1}^n w_i [1 - G(a_i - w_i - \sigma)]) \quad (\text{A.25})$$

The first order condition for worker i 's wage is,

$$(a_i - \mu w_i) g(a_i - w_i - \sigma) - \mu [1 - G(a_i - w_i - \sigma)] = 0 \quad (\text{A.26})$$

The optimal solution to the constrained optimization problem for worker i will be the same as the unconstrained problem if and only if the shadow price (μ) is one. The shadow price (μ) has an important meaning in this case. It is the increase in expected output for corresponding to a one dollar increase in the budget. Therefore if the firm could, it would set the budget such that this shadow price would be unity. In order for this to be a saddle point, the total ability of the worker under the manager must be $\frac{K}{[1-G(x^*-\sigma)]} - nx^*$. However since the budget is set prior to the observation of worker ability this happens with probability zero. When the combined ability stock is too high or too low the constrained wages are different from the optimal unconstrained wages. Therefore ex post expected profits are below optimum. ■

Proof. Of 2 The FOC for a_i can be written as

$$\left(\frac{1}{\mu} - 1\right) a_i + a_i - w_i = \frac{1}{H(a_i - w_i - \sigma)} = 0 \quad (\text{A.27})$$

If $\mu > 1$, this implies that there is too much ability relative to the budget. There will be a greater difference between ability and wages are taken from higher ability workers. As a result, turnover is greater for higher ability workers. When $\mu < 1$ the budget is high compared to average ability. More of the excess is given to the higher ability workers. Turnover is lower for the higher ability workers. ■

APPENDIX B
TURNOVER AND PROMOTION

The following shows some of the properties of the problem that we will be using repeatedly in our analysis. We will now show these in the most general terms so that we will easily use them to prove the various propositions within the paper.

Proposition 22 *Suppose that G is a CDF with a continuous, differentiable density g on its support. Where the supremum of the support of G is $\bar{x} > \varepsilon$. Let the hazard rate $H = \frac{g}{1-G}$ be non decreasing. Given the following maximization problem and solution*

$$V(\varepsilon) = \max_x x(1 - G(x + \varepsilon)) \tag{B.1}$$

The following is true.

1. *The solution to the problem $x(\varepsilon)$ is single valued, continuously differentiable with derivative $x'(\varepsilon) = \frac{-x(\varepsilon)H'(x(\varepsilon)+\varepsilon)}{H(x(\varepsilon)+\varepsilon)+x(\varepsilon)H'(x(\varepsilon)+\varepsilon)} \in (-1, 0]$*
2. *$V(\varepsilon)$ is continuously differentiable, decreasing with derivative $V'(\varepsilon) = (1-G(x(\varepsilon)+\varepsilon))$*

Proof. Of Proposition 22. 1. Since the hazard rate is non decreasing, the objective function is quasiconcave and the first order condition $xH(x + \varepsilon) = 1$ is sufficient. Also since H is continuous and non decreasing there will always exist an $x > 0$ that solves this equation. Let us label the solution $x(\varepsilon)$. By the implicit function theorem $x'(\varepsilon) = \frac{-x(\varepsilon)H'(x(\varepsilon)+\varepsilon)}{H(x(\varepsilon)+\varepsilon)+x(\varepsilon)H'(x(\varepsilon)+\varepsilon)} \in (-1, 0]$ Since non-decreasing H implies $H' \geq 0$. Part 2. follows directly from the envelope theorem. $V'(\varepsilon) = (1 - G(x(\varepsilon) + \varepsilon))$ ■

Proof. Of Proposition 12

1) From Proposition 22 $\frac{dx^*}{d\sigma} = \frac{x^*H'(x^*-\sigma)}{H(x^*-\sigma)+x^*H'(x^*-\sigma)} \in [0, 1)$.

2) This follows directly from 1. Since x^* increases with σ at a rate less than one, this implies that as σ increases $x^* - \sigma$ decreases and therefore turnover $[1 - G(x^* - \sigma)]$ decreases as well.

3) Also from Proposition 22 $\frac{dV^*}{d\sigma} = [1 - G(x^* - \sigma)]$. We know from 2) that it decreases as σ increases.

4) This follows directly from the increasing V^* in part 3) ■

Notation 3 let $\pi_{i,t}, \bar{\pi}_{i,t} \in \{0, 1\}$ represent the training decision for the current and outside firms respectively. Since training costs are born once, we only consider the decision to train a previously untrained worker. Let $\pi_{i,t} = 1$ if worker i 's current firm decides to train him at time t else $\pi_{i,t} = 0$. Let $\bar{\pi}_{i,t} = 1$ if an outside firm trains worker i given that (untrained) worker i enters the outside market at time t else $\bar{\pi}_{i,t} = 0$. For simplicity, we assume firms offer promotion when indifferent.

Definition 5 Given worker i , equilibrium is a series of promotion decisions for current and outside firms, $\{\pi_{i,t}\}, \{\bar{\pi}_{i,t}\}$ current wages for un promoted and promoted workers $\{w_{i,1,t}\}, \{w_{i,2,t}\}$, outside wage offers with no promotion $\{\bar{w}_{i,1,t}\}$, outside firm wage offers with promotion $\{\bar{w}_{i,2,t}\}$, outside firm wage offers if already promoted $\{\hat{w}_{i,2,t}\}$, and expected profits for the current firm on a worker that was not promoted, and promoted, $\{V_{i,1,t}\}, \{V_{i,2,t}\}$, ex. post expected future utilities of a promoted and un promoted worker $\{U_{i,1,t}\}, \{U_{i,2,t}\}$ such that for all $t \geq 1$ Conditions 6, 4, 5, and 7 are satisfied.

Condition 4 (*Zero Profit for Outside Market*)

- 1) $\hat{w}_{i,2,t} = y_{i,2,t} + \delta V_{i,2,t+1}$
- 2) $\bar{w}_{i,2,t} = y_{i,2,t} + \delta V_{i,2,t+1} - b$,
- 3) $\bar{w}_{i,1,t} = y_{i,1,t} + \delta V_{i,1,t+1}$,
- 4) $\bar{\pi}_{i,t} = \begin{cases} 1 & \text{if } \bar{w}_{i,1,t} + \delta U_{i,1,t+1} \leq \bar{w}_{i,2,t} + \delta U_{i,2,t+1} \\ 0 & \text{if } \bar{w}_{i,1,t} + \delta U_{i,1,t+1} > \bar{w}_{i,2,t} + \delta U_{i,2,t+1} \end{cases}$

Condition 5 (*Expected Profits for Current Firm*)

- 1) $V_{i,2,t} = \max_w (y_{i,2,t} + \delta V_{i,2,t+1} - w) [1 - G(\hat{w}_{i,2,t} - w - \sigma)]$
- 2) $V_{i,1,t} = \max \begin{cases} V_{i,2,t} - b \\ \max_w (y_{i,1,t} + \delta V_{i,1,t+1} - w) [1 - G(u_{i,t}^o - w - \sigma)] \end{cases}$

Once the worker has been promoted the definition of V is analogous to the one job model. For a worker i who has yet to be trained, the firm can train him at the beginning of period t or not. Once worker i is trained, the current firm treats training costs as sunk. The wage decision will be the same as if the worker was promoted in an earlier period. Therefore, If he trains worker i in period t , the firm's payoff is $-b$ plus the same payoff as if the worker was trained prior to t . If the current firm doesn't train at t , he chooses a potentially different wage offer.

Condition 6 (*Profit Maximization for Current Firm*)

- 1) $w_{i,1,t} \in \arg \max_w (y_{i,1,t} + \delta V_{i,1,t+1} - w) [1 - G(u_{i,t}^o - w - \sigma)]$
- 2) $w_{i,2,t} \in \arg \max_w (y_{i,2,t} + \delta V_{i,2,t+1} - w) [1 - G(\hat{w}_{i,2,t} - w - \sigma)]$
- 3) $\pi_{i,t} = \begin{cases} 0 & \text{if } V_{i,1,t} > V_{i,2,t} - b \\ 1 & \text{else} \end{cases}$

where

$$u_{i,t}^o = \max \{ \bar{w}_{i,1,t} + \delta(U_{i,1,t+1} - U_{i,1,t+1}), \bar{w}_{i,2,t} + \delta(U_{i,2,t+1} - U_{i,1,t+1}) \} \quad (\text{B.2})$$

$u_{i,t}^o$ is untrained worker i 's expected wages and additional future utility for entering the outside market.

Condition 7 (*Expected Utility*)

$$1) U_{i,2,t} = \int_{-\infty}^{\infty} \max\{w_{i,2,t} + \delta U_{i,2,t+1} + \lambda, \hat{w}_{i,2,t} + \delta U_{i,2,t+1} - \sigma\} g(\lambda) d\lambda$$

$$2) U_{i,1,t} = \begin{cases} U_{i,2,t} & \text{if } \pi_{i,t} = 1 \\ \int_{-\infty}^{\infty} \max\{w_{i,1,t} + \delta U_{i,1,t+1} + \lambda, u_{i,t}^o + \delta U_{i,1,t+1} - \sigma\} g(\lambda) d\lambda & \text{if } \pi_{i,t} = 0 \end{cases}$$

Proof. of proposition 13

We know from proposition 11 that the equilibrium conditions are satisfied for period in which the worker has already been promoted. We now show that both the current and outside firms will offer promotion after some finite time. Once we know this, we use backward induction to solve the rest of the game to show that the solution exists and is unique.

1), and 2) follow directly from identical analysis as proposition 11.

3) If $\bar{\pi}_{i,t} = 0$, the outside firms are not offering promotion at the wage, $\bar{w}_{i,1,t} = y_{i,1,t} + \delta V_{i,1,t+1}$. If the current firm does not offer promotion, his expected profits are, $\max_w (y_{i,1,t} + \delta V_{i,1,t+1} - w) [1 - G(y_{i,1,t} + \delta V_{i,1,t+1} - \sigma)] = V^*$. Solving this problem we see that the current firm's wage offer will be, $w_{i,1,t} = y_{i,1,t} + \delta V_{i,1,t+1} - x^*$. If it offers promotion, its expected profits are $V^* - b$. Therefore if the current firm expects the outside firms will not promote the worker, it will not offer promotion either.

4) In order to show this we must first know that in any equilibrium, $V_{i,1,t}$ is bounded above by V^* . Also we must know that future worker expected utility is greater for a promoted worker, when the worker is more productive at the higher job. Once we know this, $y_{i,2,t} \geq y_{i,1,t} \Rightarrow U_{i,2,t+1} \geq U_{i,1,t+1}$ $y_{i,2,t} - b \geq y_{i,1,t} \Rightarrow$

$y_{i,2,t} - b + \delta V^* + \delta U_{i,2,t+1} \geq y_{i,1,t} + \delta V_{i,1,t} + \delta U_{i,1,t+1}$. Therefore, $\bar{w}_{i,2,t} + \delta U_{i,2,t+1} \geq \bar{w}_{i,1,t} + \delta U_{i,1,t+1}$. By Condition 4 the outside firms offer promotion.

We now prove that the sufficient conditions used above satisfied in any equilibrium. We first start with the assumption on V . If the current firm offers promotion $V_{i,1,t} = V^* - b < V^*$. If the current firm doesn't offer promotion, $V_{i,1,t} = \max_w (y_{i,1,t} + \delta V_{i,1,t+1} - w) [1 - G(u_{i,t}^o - w - \sigma)]$. The value of $V_{i,1,t}$ is decreasing with $u_{i,t}^o$. Since $u_{i,t}^o$ is bounded below by $y_{i,1,t} + \delta V_{i,1,t+1}$, $V_{i,1,t}$ must be bounded above by V^* by proposition 22.

We will now prove that $y_{i,2,t} > y_{i,1,t} \Rightarrow U_{i,2,t+1} \geq U_{i,1,t+1}$. The expected utility If nobody offers promotion can be written as,

$$U_{i,1,t} = y_{i,1,t} + \delta V_{i,1,t+1} + \delta U_{i,1,t+1} + \int_{-\infty}^{\infty} \max\{\lambda - x^*, -\sigma\} g(\lambda) d\lambda. \quad (\text{B.3})$$

If the worker has already been promoted expected utility can be written as

$$U_{i,2,t} = y_{i,2,t} + \delta V^* + \delta U_{i,2,t+1} + \int_{-\infty}^{\infty} \max\{\lambda - x^*, -\sigma\} g(\lambda) d\lambda. \quad (\text{B.4})$$

To simplify the analysis for the worker who is only promoted by outside firms, I introduce the following notation.

$$\phi_{i,t} = y_{i,2,t} - y_{i,1,t} + \delta (V^* - V_{i,1,t+1}) + \delta (U_{i,2,t+1} - U_{i,1,t+1}) - b. \quad (\text{B.5})$$

$\phi_{i,t}$ Can be interpreted as the difference between the non current random shock component of expected utility for a promoted vs. un promoted worker. (Note, that outside firms will offer promotion whenever $\phi_{i,t} \geq 0$.) The expected utility of a worker that will be offered promotion from only the outside market at time t is,

$$U_{i,1,t} = y_{i,2,t} + \delta V^* + \delta U_{i,2,t+1} + \quad (\text{B.6})$$

$$\int_{-\infty}^{\infty} \max\{-\phi_{i,t} + \lambda - x(\phi_{i,t}), -\sigma\} g(\lambda) d\lambda$$

$$\text{Where } x(\phi_{i,t}) = \arg \max_x [1 - G(\phi_{i,t} + x - \sigma)] \quad (\text{B.7})$$

The difference in expected future utilities between a promoted vs. un promoted worker becomes very important in employment decision for the worker. We will now show that the difference between expected utility for a promoted vs un promoted worker will be non negative Whenever $y_{i,2,t} \geq y_{i,1,t}$.

Case 1. (both current and outside firms offer promotion) If both the current and outside firms offer promotion at time t then for a worker at the beginning of time t his expected utility (before observing taste shock) will be the same whether or not he has previously been promoted. i.e. $U_{i,1,t} = U_{i,2,t}$.

Case 2. (only outside market offers promotion) If at time t only the outside market offers promotion then $\phi_{i,t} \geq 0$ and the difference

$$\begin{aligned} & \int_{-\infty}^{\infty} \max\{\lambda - x^*, -\sigma\}g(\lambda)d\lambda \\ & - \int_{-\infty}^{\infty} \max\{-\phi_{i,t} + \lambda - x(\phi_{i,t}), -\sigma\}g(\lambda)d\lambda \\ & = U_{i,2,t} - U_{i,1,t} \geq 0. \end{aligned} \tag{B.8}$$

Since $x'(\cdot) > -1$ (from Proposition 22) and $\phi_{i,t} \geq 0$ implies $\phi_{i,t} + x(\phi_{i,t}) \geq 0 + x(0) = x^*$

Case 3 (neither current nor outside firms offer promotion) Proof by contradiction. Suppose that at t neither the current or the outside firm offers promotion to an un promoted worker i . Also assume, $y_{i,2,t} > y_{i,1,t}$ and $U_{i,2,t} < U_{i,1,t}$. If no firms offer the worker promotion at time t then the difference in current and outside utility from (B.3) and (B.4) is

$$U_{i,2,t} - U_{i,1,t} = y_{i,2,t} - y_{i,1,t} + \delta(V^* - V_{i,1,t+1}) + \delta(U_{i,2,t+1} - U_{i,1,t+1}) \tag{B.9}$$

$y_{i,2,t} > y_{i,1,t}$, $V^* \geq V_{i,1,t+1}$, and $U_{i,2,t} < U_{i,1,t}$ implies that $U_{i,2,t+1} < U_{i,1,t+1}$. Because we know that $y_{i,2,t+1} > y_{i,1,t+1}$, neither the inside nor the outside firm will offer

promotion at time $t + 1$ (by cases 1 and 2). By forward induction it must be true that the employee will never be offered promotion in the future. Therefore we solve for what the worker's lifetime expected utility if not promoted at time t .

$$U_{i,1,t} = \sum_{k=0}^{\infty} \delta^k y_{i,1,t+k} + \frac{\delta V^*}{1-\delta} + \frac{\delta \int_{-\infty}^{\infty} \max\{\lambda - x^*, -\sigma\} g(\lambda) d\lambda}{1-\delta}. \quad (\text{B.10})$$

And if he is promoted at time t ,

$$U_{i,2,t} = \sum_{k=0}^{\infty} \delta^k y_{i,2,t+k} + \frac{\delta V^*}{1-\delta} + \frac{\delta \int_{-\infty}^{\infty} \max\{\lambda - x^*, -\sigma\} g(\lambda) d\lambda}{1-\delta}. \quad (\text{B.11})$$

Since $y_{i,2,t} - y_{i,1,t} \geq 0$ and y_2 is growing faster than y_1 , $U_{i,2,t} - U_{i,1,t} = \sum_{k=0}^{\infty} \delta^k (y_{i,2,t+k} - y_{i,1,t+k}) > 0$. This is a contradiction to the assumption that $U_{i,2,t} < U_{i,1,t}$. Therefore it must be the case that $U_{i,2,t} \geq U_{i,1,t}$ whenever $y_{i,2,t} \geq y_{i,1,t}$. Therefore the outside firm will offer promotion whenever $y_{i,2,t} - b \geq y_{i,1,t}$.

5) $y_{i,2,t_I} - y_{i,1,t_I} > (1-\delta)b + \hat{\phi}^b$ implies that outside firms offer promotion since $\hat{\phi}^b > b$. Therefore

$$\begin{aligned} \phi_{i,t} &= y_{i,2,t_I} - y_{i,1,t_I} + \delta(V^* - V_{i,1,t_I+1}) \\ &\quad + \delta(U_{i,2,t_I+1} - U_{i,1,t_I+1}) - b \\ &\geq \hat{\phi}^b. \end{aligned} \quad (\text{B.12})$$

Since $U_{i,2,t_I+1} \geq U_{i,1,t_I+1}$ and $V^* - V_{i,1,t_I+1} \geq b$, $\phi_{i,t} \geq \hat{\phi}^b$ implies that

$$\max x [1 - G(\phi_{i,t} - x - \sigma)] \leq \max x [1 - G(\hat{\phi}^b - x - \sigma)] = V^* - b,$$

by definition of $\hat{\phi}^b$. Therefore the current firm will promote the worker since he weakly prefers to do so.

6) $y_{i,2,\bar{t}+1} - y_{i,1,\bar{t}+1} \geq (1-\delta)b + \hat{\phi}^b \Rightarrow$ worker i will be promoted at $\bar{t} + 1$. Therefore $U_{i,2,\bar{t}+1} = U_{i,1,\bar{t}+1}$ and $y_{i,2,\bar{t}} - y_{i,1,\bar{t}} < (1-\delta)b + \hat{\phi}^b$ implies,

$$\phi_{i,t} = y_{i,2,\bar{t}} - y_{i,1,\bar{t}} + \delta(V^* - V_{i,1,t+1}) - b < \hat{\phi}^b. \quad (\text{B.13})$$

Therefore the inside firm will not offer promotion at \bar{t} . If $y_{i,2,\bar{t}} - y_{i,1,\bar{t}} > b$ then $\phi_{i,t} > 0$. Outside firms offer promotion. Turnover rates are higher than after promotion or before worker i is offered promotion at all. That is,

$$1 - G(\phi_{i,\bar{t}} + x(\phi_{i,\bar{t}}) - \sigma) > 1 - G(x^* - \sigma). \quad (\text{B.14})$$

■

Condition 8

$$\delta \leq \frac{(c_2 - c_1)a_i}{\int_{-\infty}^{\infty} \max\{\lambda - x^*, -\sigma\}g(\lambda)d\lambda - \int_{-\infty}^{\infty} \max\{-\hat{\phi}^b + \lambda - x(\hat{\phi}^b), -\sigma\}g(\lambda)d\lambda}. \quad (\text{B.15})$$

Proof. Of Proposition 14. We start by showing that $\phi_{i,t}$ is increasing each period when $t \leq T_i$. We will do this using an induction proof. All other results follow directly from the increasing nature of $\phi_{i,t}$. Let $T_i \in \mathbb{R}$ Solve $y_{i,2,T_i} - y_{i,1,T_i} = (1 - \delta)b + \hat{\phi}^b$. Let $t^* = \max\{t | t < T_i\}$. Promotion will be offered by the current firm at time $t^* + 1$. Therefore $V^* - V_{i,1,t^*+1} = b$, and $U_{i,2,t^*+1} = U_{i,1,t^*+1}$

From Condition 8,

$$\delta \leq \frac{(c_2 - c_1)a_i}{\int_{-\infty}^{\infty} \max\{\lambda - x^*, -\sigma\}g(\lambda)d\lambda - \int_{-\infty}^{\infty} \max\{-\hat{\phi}^b + \lambda - x(\hat{\phi}^b), -\sigma\}g(\lambda)d\lambda} \quad (\text{B.16})$$

$$< \frac{(c_2 - c_1)a + \delta(V_{i,1,t^*} - (V^* - b))}{\int_{-\infty}^{\infty} \max\{\lambda - x^*, -\sigma\}g(\lambda)d\lambda - \int_{-\infty}^{\infty} \max\{-\phi_{i,t^*} + \lambda - x(\phi_{i,t^*}), -\sigma\}g(\lambda)d\lambda} \quad (\text{B.17})$$

$$= \frac{(c_2 - c_1)a + \delta(V_{i,1,t^*} - (V^* - b))}{U_{i,2,t^*} - U_{i,1,t^*}}. \quad (\text{B.18})$$

Since $\phi_{i,t^*} < \hat{\phi}^b$ and $V_{i,1,t^*} \geq V^* - b$. By the linearity assumption $(c_2 - c_1)a = y_{i,2,t^*} - y_{i,2,t^*-1} - (y_{i,1,t^*} - y_{i,1,t^*-1})$. Rearranging the inequality we get,

$$\phi_{i,t^*} = y_{i,2,t^*} - y_{i,1,t^*} + \delta b - b \quad (\text{B.19})$$

$$> y_{i,2,t^*-1} - y_{i,1,t^*-1} + \delta(V^* - V_{i,1,t^*}) + \delta(U_{i,2,t^*} - U_{i,1,t^*}) - b \quad (\text{B.20})$$

$$= \phi_{i,t^*-1}. \quad (\text{B.21})$$

Now that we have shown $\phi_{i,t^*} > \phi_{i,t^*-1}$, we will perform the induction step. Suppose that $\hat{\phi}^b \geq \phi_{i,t} \geq \phi_{i,t-1}$ for some $t \leq t^*$. Since $\max x [1 - G(\phi - x - \sigma)]$ is decreasing in ϕ then $V_{i,1,t} < V_{i,1,t-1}$. Since $x'(\cdot) > -1 \Rightarrow -\max\{-\phi_{i,t} + \lambda - x(\phi_{i,t}), -\sigma\}$ is increasing in $\phi_{i,t}$. Therefore $\phi_{i,t} \geq \phi_{i,t-1} \Rightarrow U_{i,2,t} - U_{i,1,t} \geq U_{i,2,t-1} - U_{i,1,t-1} \Rightarrow$

$$\begin{aligned}
\phi_{i,t-1} &= y_{i,2,t-1} - y_{i,1,t-1} + \delta(V^* - V_{i,1,t}) + \delta(U_{i,2,t} - U_{i,1,t}) - b \quad (\text{B.22}) \\
&> y_{i,2,t-2} - y_{i,1,t-2} + \delta(V^* - V_{i,1,t-1}) \\
&\quad + \delta(U_{i,2,t-1} - U_{i,1,t-1}) - b \\
&= \phi_{i,t-2}
\end{aligned}$$

Since we know that $\phi_{i,t^*} > \phi_{i,t^*-1}$ we also know that $\phi_{i,t}$ is increasing, and expected profits $V_{i,1,t}$ are decreasing whenever $t \leq T_i$

1) Since we know that $\phi_{i,t}$ is increasing until time period T_i we know that $\phi_{i,t} < \hat{\phi}^b$ for all $t < T_i$. Therefore the current firm will never offer promotion before T_i

2) We know that $\phi_{i,t}$ is increasing for $t < T_i$ and ≥ 0 when $y_{i,2,t} - y_{i,1,t} \geq b$. There will be a time \bar{T}_i where $\phi_{i,t} < 0$ if and only if $t < \bar{T}_i$. That is, the outside market will always offer promotion after \bar{T}_i but never before.

3) For any period $t' \in (\bar{T}_i, T_i)$ the current firm will never offer promotion before this time. Since each period turnover rates are less than 1 there is a positive probability that the worker will be un promoted at time t' . $t' > \bar{T}_i \Rightarrow \phi_{i,t} > 0 \Rightarrow$ the turnover rate $(1 - G(\phi_{i,t} + x(\phi_{i,t}) - \sigma))$ will be greater than after promotion, or when no firms were offering promotion.

4) This follows directly from the increasing nature of $\phi_{i,t}$ ■

Proof. of Corollary 3 Let t^* be the last period before promotion is offered from within the firm. By assumption the outside market is offering promotion at t^* that is

$$\begin{aligned}\phi_{i,t^*} &= y_{i,2,t^*} - y_{i,1,t^*} + \delta(V^* - V_{i,1,t^*+1}) + \delta(U_{i,2,t^*+1} - U_{i,1,t^*+1}) - b \text{ (B.23)} \\ &= y_{i,2,t^*} - y_{i,1,t^*} - (1 - \delta)b > 0.\end{aligned}$$

Let

$$x(\phi_{i,t^*}) = y_{i,1,t^*} + \delta V_{i,1,t^*+1} - w_{i,1,t^*}. \quad \text{(B.24)}$$

Since $x'(\phi) > -1$ and $\phi_{i,t^*} > 0$ we know that

$$\phi_{i,t^*} - x(\phi_{i,t^*}) > x(0) = x^*. \quad \text{(B.25)}$$

Substituting in $x^* = y_{i,2,t^*+1} + \delta V^* - w_{i,2,t^*+1}$, (B.23) and (B.24) we get

$$y_{i,2,t^*} + \delta(U_{i,2,t^*+1} - U_{i,1,t^*+1}) - w_{i,1,t^*} - b \geq y_{i,2,t^*+1} - w_{i,2,t^*+1}. \quad \text{(B.26)}$$

Therefore,

$$w_{i,2,t^*+1} - w_{i,1,t^*} \geq y_{i,2,t^*+1} - y_{i,2,t^*} + b = c_2 a_i + b. \quad \text{(B.27)}$$

Wage growth at promotion must be greater than $a_i c_2 + b$. And as we have seen in the simple model, wage growth after promotion must be exactly $a_i c_2$. Therefore we have shown that wage growth after promotion must be smaller than before promotion.

Now we will show that wage increases the period before promotion must be less than c_2 . Under condition 8 we have shown that $\phi_{i,t}$ is increasing until promotion. We need to look at wage increases for both the conditions if the outside market offers promotion at $t^* - 1$ and when it doesn't.

Case 1. The outside market is offering promotion at both t^* and $t^* - 1$. Suppose $\phi_{i,t^*} \geq \phi_{i,t^*-1}$, $x'(\phi) > -1$ implies,

$$\phi_{i,t^*} + x(\phi_{i,t^*}) \geq \phi_{i,t^*-1} + x(\phi_{i,t^*-1}). \quad (\text{B.28})$$

Substituting in (B.23) and (B.24) for both t^* and $t^* - 1$ implies

$$y_{i,2,t^*} + \delta(U_{i,2,t^*+1} - U_{i,1,t^*+1}) - w_{i,1,t^*} \geq y_{i,2,t^*-1} + \delta(U_{i,2,t^*} - U_{i,1,t^*}) - w_{i,1,t^*-1} \quad (\text{B.29})$$

since promotion is assured at $t^* + 1$, $U_{i,2,t^*+1} = U_{i,1,t^*+1}$ and this becomes

$$y_{i,2,t^*} - y_{i,2,t^*-1} - \delta(U_{i,2,t^*} - U_{i,1,t^*}) \geq w_{i,1,t^*} - w_{i,1,t^*-1} \quad (\text{B.30})$$

Wage growth before promotion is bounded above $y_{i,2,t^*} - y_{i,2,t^*-1} = c_2 a_i$.

Suppose instead that $\phi_{i,t^*} < \phi_{i,t^*-1}$, $x'(\phi) \leq 0 \Rightarrow x(\phi_{i,t^*}) \geq x(\phi_{i,t^*-1}) \Rightarrow y_{i,1,t^*} + \delta V_{i,1,t^*+1} - w_{i,1,t^*} \geq y_{i,1,t^*-1} + \delta V_{i,1,t^*} - w_{i,1,t^*-1}$

$\Rightarrow y_{i,1,t^*} - y_{i,1,t^*-1} + \delta V_{i,1,t^*+1} - \delta V_{i,1,t^*} \geq w_{i,1,t^*} - w_{i,1,t^*-1}$ Wage growth is bounded above by $y_{i,1,t^*} - y_{i,1,t^*-1} = c_1 a_i < c_2 a_i$.

Case 2. The outside market is offering promotion at t^* but not $t^* - 1$. If the outside market does not offer promotion at $t^* - 1$ we know that

$$w_{i,1,t^*-1} = y_{i,1,t^*-1} + \delta V_{i,1,t^*} - x^*. \quad (\text{B.31})$$

Since the firm will offer promotion at $t^* + 1$,

$$w_{i,2,t^*+1} = y_{i,2,t^*+1} + \delta V^* - x^*. \quad (\text{B.32})$$

From (B.27),

$$w_{i,2,t^*+1} - w_{i,1,t^*} \geq c_2 a_i + b \quad (\text{B.33})$$

$$\Rightarrow w_{i,2,t^*+1} - c_2 a_i - b - w_{i,1,t^*-1} \geq w_{i,1,t^*} - w_{i,1,t^*-1} \quad (\text{B.34})$$

Substituting (B.31) and (B.32) into the left hand side and rearranging the inequality gives us,

$$c_2 a_i > c_2 a_i + (y_{i,2,t^*-1} - y_{i,1,t^*-1}) + \delta (V^* - V_{i,1,t^*}) - b \geq w_{i,1,t^*} - w_{i,1,t^*-1}. \quad (\text{B.35})$$

Since the outside firms offer promotion t^* but not at $t^* - 1 \Rightarrow (y_{i,2,t^*-1} - y_{i,1,t^*-1}) + \delta (V^* - V_{i,1,t^*}) - b < 0$. Therefore the wage increase before promotion is less than the wage increase after promotion. ■

Proof. of Corollary 4 Given $w_{i,1,t^*} = y_{i,1,t^*} + \delta (V^* - b) - x(\phi_{i,t^*})$ and $w_{i,2,t^*+1} = y_{i,2,t^*+1} + \delta V^* - x^*$,

$$w_{i,2,t^*+1} - w_{i,1,t^*} = a_i c_2 + \phi_{i,t^*} + b + x(\phi_{i,t^*}) - x^*. \quad (\text{B.36})$$

$$\text{Where } \phi_{i,t^*} = y_{i,2,t^*} - y_{i,1,t^*} - (1 - \delta)b \quad (\text{B.37})$$

Since ϕ_{i,t^*} is increasing in the difference in productivities at the two job assignments before promotion, and $\phi_{i,t^*} + x(\phi_{i,t^*})$ is increasing in ϕ_{i,t^*} . The wage premium at promotion is also increasing in the difference in productivities at the two job assignments before promotion. ■

Proof. Of Corollary 5. Recall that $\hat{\phi}^b$ solves the equality,

$$\max_x \left[1 - G(x + \hat{\phi}^b - \sigma) \right] = \max_z \left[1 - G(z - \sigma) \right] - b.$$

Both maximization problems have a unique solution and they are continuous and differentiable in both $\hat{\phi}^b$ and σ . Using the implicit function theorem to find the effect that σ has on $\hat{\phi}^b$ we get

$$\frac{\partial \hat{\phi}^b}{\partial \sigma} = \frac{[1 - G(x^* - \sigma)] - [1 - G(x(\hat{\phi}^b) + \hat{\phi}^b - \sigma)]}{[1 - G(x(\hat{\phi}^b) + \hat{\phi}^b - \sigma)]} < 0 \quad (\text{B.38})$$

We know that $0 < b < \hat{\phi}^b$ therefore by Proposition 22, turnover will be less for the promoted worker than for the un promoted worker $(1 - G(x^* - \sigma)) >$

$1 - G(x(\hat{\phi}^b) + \hat{\phi}^b - \sigma)$ This implies that, given the worker's switching cost, the current firm will offer promotion when $y_{i,2,t} - y_{i,1,t} > \hat{\phi}^b + (1 - \delta)b$. This threshold of promotion $\hat{\phi}^b$ is lower for a worker with a higher switching cost. ■

APPENDIX C

EMPLOYEE BONDING AND TURNOVER EFFICIENCY

Proof. of Proposition 17 The manager chooses a wage to maximize second period expected profits,

$$V_{i,f}(b) = \max_w (a_i - w) [1 - G(\bar{w}_i - w - \sigma - b)]. \quad (\text{C.1})$$

As in the previous sections the outside market offers a wage equal to the worker's marginal product. His optimal decision is now

$$w_{i,f}(b) = \arg \max_w (a_i - w) [1 - G(a_i - w - \sigma - b)]. \quad (\text{C.2})$$

rewriting the optimal ex post rent as a function of the bond we get $x(b) = a_i - w_{i,f}(b)$. Since the objective function is quasiconcave this must solve,

$$x_i(b) = \frac{1}{H(x_i(b) - \sigma - b)}. \quad (\text{C.3})$$

From the implicit function theorem the derivative of $x_i(b)$ with respect to b is,

$$x'_i(b) = \frac{x_i(b) H'(x_i(b) - \sigma - b)}{H(x_i(b) - \sigma - b) + x_i(b) H'(x_i(b) - \sigma - b)} \in (0, 1). \quad (\text{C.4})$$

Since $x_i(b)$ is increasing with b at a slower rate than 1 this implies that the turnover rate $G(x_i(b) - \sigma - b)$ is decreasing with the bond b . ■

Proof. of Propostion 18. Since there is Bertrand like competition in the first period, the equilibrium bond size and first period wage maximize expected efficiency. Second period optimal turnover happens when the wage is equal to marginal product minus the bond. In the second period, managers choose wages to solve the following first order condition $a_i - w_i = \frac{1-G(a_i-w_i-b-\sigma)}{g(a_i-w_i-b-\sigma)}$. Therefore a bond equal to $b^* = \frac{1-G(-\sigma)}{g(-\sigma)}$ will induce the optimal turnover rate of $G(-\sigma)$ in the second period.

The optimal bond size is unique since the turnover rate is decreasing with the size of the bond. ■

Proof. of Proposition 20. The manager chooses a wage to maximize second period expected profits,

$$V_{i,t}(b) = \max_w (a_i + b - w) [1 - G(\bar{w}_i - w - \sigma - b)]. \quad (\text{C.5})$$

As in the previous sections the outside market offers a wage equal to the worker's marginal product. His optimal decision is now

$$w_{i,t}(b) = \arg \max_w (a_i + b - w) [1 - G(a_i - w - \sigma)]. \quad (\text{C.6})$$

rewriting the optimal ex post rent as a function of the bond we get $x_i(b) = a_i + b - w_{i,t}(b)$. Since the objective function is quasiconcave this must solve,

$$x_i(b) = \frac{1}{H(x_i(b) - \sigma - b)}. \quad (\text{C.7})$$

From the implicit function theorem the derivative of $x_i(b)$ with respect to b is,

$$x'_i(b) = \frac{x_i(b) H'(x_i(b) - \sigma - b)}{H(x_i(b) - \sigma - b) + x_i(b) H'(x_i(b) - \sigma - b)} \in (0, 1). \quad (\text{C.8})$$

Since $x_i(b)$ is increasing with b at a slower rate than 1 this implies that the turnover rate $G(x_i(b) - \sigma - b)$ is decreasing with the training b . Recall that the wage $w_{i,t}(b) = a_i + b - x(b)$. Training results in an increase in the wage that is smaller than the training cost. ■

Proof. of Proposition 21. The manager chooses a wage to maximize second period expected profits,

$$V_{i,t}(K) = \max_w (a_i + K - w) [1 - G(\bar{w}_i - w - \sigma)]. \quad (\text{C.9})$$

As in the previous sections the outside market offers a wage equal to the worker's marginal product. His optimal decision is now

$$w_{i,t}(K) = \arg \max_w (a_i + K - w) [1 - G(a_i - w - \sigma)]. \quad (\text{C.10})$$

rewriting the optimal ex post rent as a function of the bond we get $x_i(b) = a_i + b - w_{i,f}(b)$. Since the objective function is quasiconcave this must solve,

$$x_i(K) = \frac{1}{H(x_i(K) - \sigma - K)}. \quad (\text{C.11})$$

From the implicit function theorem the derivative of $x_i(b)$ with respect to b is,

$$x'_i(K) = \frac{x_i(K) H'(x_i(K) - \sigma - K)}{H(x_i(K) - \sigma - K) + x_i(K) H'(x_i(K) - \sigma - K)} \in (0, 1). \quad (\text{C.12})$$

Since $x_i(K)$ is increasing with K at a slower rate than 1 this implies that the turnover rate $G(x_i(K) - \sigma - K)$ is decreasing with the training K . Recall that the wage $w_{i,t}(K) = a_i + K - x(K)$. Training results in an increase in the wage that is smaller than the training cost. ■

BIBLIOGRAPHY

Acemoglu, D. and J. S. Pischke. 1999. "The Structure of Wages and Investment in General Training" *Journal of Political Economy*, 107(3): 539.

Acemoglu, D. and J. S. Pischke. 1998. "Why do Firms Train? Theory and Evidence" *The Quarterly Journal of Economics*, 113(1): 79-119.

Allen, S. G., R. L. Clark, and A. A. McDermid. 1993. "Pensions, Bonding, and Lifetime Jobs" *Journal of Human Resources*, 28(3): 463-481.

Arrow, K. J. 1962. "The Economic Implications of Learning by Doing" *The Review of Economic Studies*, 29(3): 155-173.

Baker, G. P., M. C. Jensen, and K. J. Murphy. 1988. "Compensation and Incentives: Practice Vs. Theory" *The Journal of Finance*, 43(3): 593-616.

Baker, George, Michael Gibbs, and Bengt Holmstrom. 1994. "The Wage Policy of a Firm" *The Quarterly Journal of Economics*, 109(4): 921-955.

Bar-Isaac, H., I. Jewitt, and C. Leaver. 2008. "Information and Human Capital Management" .

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

Becker, G. S. 1962. "Investment in Human Capital: A Theoretical Analysis" *The Journal of Political Economy*, 70(5): 9-49.

Bernhardt, D. 1995. "Strategic Promotion and Compensation" *The Review of Economic Studies*, 62(2): 315-339.

Brown, James 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.

Buchmueller, T. C. and R. G. Valletta. 1995. "Effects of Employer-Provided Health Insurance on Worker Mobility, the" *Indus.& Lab.Rel.Rev.*, 49: 439.

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

- Cappelli, P. 2004. "Why do Employers Pay for College?" *Journal of Econometrics*, 121(1-2): 213-241.
- Carmichael, L. 1985. "Can Unemployment be Involuntary? Comment" *The American Economic Review*, 75(5): 1213-1214.
- Carson, P. P., K. D. Carson, R. W. Griffeth, and R. P. Steel. 1994. "Promotion and Employee Turnover: Critique, Meta-Analysis, and Implications" *Journal of Business and Psychology*, 8(4): 455-466.
- Cascio, Wayne F. 1999. *Costing Human Resources : The Financial Impact of Behavior*, 4th ed. ed: Cincinnati, OH : South-Western College Pub.,.
- Chang, Chun and Yijiang Wang. 1996. "Human Capital Investment Under Asymmetric Information: The Pigovian Conjecture Revisited" *Journal of Labor Economics*, 14(3): 505-519.
- Costa, J. E. R. I. 1988. "Managerial Task Assignment and Promotions" *Econometrica: Journal of the Econometric Society*, 56(2): 449-466.
- Cotton, J. L. and J. M. Tuttle. 1986. "Employee Turnover: A Meta-Analysis and Review with Implications for Research" *The Academy of Management Review*, 11(1): 55-70.
- Demougin, D. and A. Siow. 1994. "Careers in Ongoing Hierarchies" *The American Economic Review*, 84(5): 1261-1277.
- DePasquale, J. A. and R. A. Lange. 1971. "Job Hopping and the MBA" *Harvard business review*, 49(6): 4-13.
- DeVaro, J. and M. Waldman. 2006. "The Signaling Role of Promotions: Further Theory and Empirical Evidence" Cornell University.
- Doeringer, P. and Piore, M. 1971. *Internal Labor Markets and Manpower Analysis*. Lexington, MA: Heath.
- Dohmen, T. J. 2004. "Performance, Seniority, and Wages: Formal Salary Systems and Individual Earnings Profiles" *Labour Economics*, 11(6): 741-763.
- Doiron, Denise J. 1995. "Lay-Offs as Signals: The Canadian Evidence" *The Canadian Journal of Economics / Revue canadienne d'Economique*, 28(4a): 899-913.

Even, William E. and David A. Macpherson. 1996. "Employer Size and Labor Turnover: The Role of Pensions" *Industrial and Labor Relations Review*, 49(4): 707-728.

Garicano, Luis. 2000. "Hierarchies and the Organization of Knowledge in Production" *The Journal of Political Economy*, 108(5): 874-904.

Gathmann, C. and U. Schönberg. 2007. "How General is Human Capital? A Task-Based Approach" , IZA Discussion Paper No. 3067.

GERHART, B. A. and G. T. MILKOVICH. 1989. "Salaries, Salary Growth, and Promotions of Men and Women in a Large, Private Firm" *Pay Equity: Empirical Inquiries*.

Ghosh, S. 2007. "Job Mobility and Careers in Firms" *Labour Economics*, 14(3): 603-621.

Gibbons, R. and L. F. Katz. 1991. "Layoffs and Lemons" *Journal of Labor Economics*: 351-380.

Gibbons, Robert and Michael Waldman. 1999. "A Theory of Wage and Promotion Dynamics Inside Firms" *The Quarterly Journal of Economics*, 114(4): 1321-1358.

Gibbons, Robert and Michael Waldman. 2004. "Task-Specific Human Capital" *The American Economic Review*, 94(2, Papers and Proceedings of the One Hundred Sixteenth Annual Meeting of the American Economic Association San Diego, CA, January 3-5, 2004): 203-207.

Gibbons, Robert and Michael Waldman. 2006. "Enriching a Theory of Wage and Promotion Dynamics Inside Firms" *Journal of Labor Economics*, 24(1): 59-107.

Gibbs, M. and W. Hendricks. 2004. "Do Formal Salary Systems really Matter?" *Industrial and Labor Relations Review*, 58(1): 71-93.

Golan, Limor. 2005. "Counteroffers and Efficiency in Labor Markets with Asymmetric Information" *Journal of Labor Economics*, 23(2): 373-393.

Golan, Limor. 2009. "Wage Signaling: A Dynamic Model of Intrafirm Bargaining and Asymmetric Learning" *International Economic Review*.

Golan, Limor. "Wage Signaling" .

- Greenwald, B. C. 1986. "Adverse Selection in the Labour Market" *The Review of Economic Studies*, 53(3): 325-347.
- Grund, Christian. 1999. "Stigma Effects of Layoffs? Evidence from German Micro-Data" *Economics Letters*, 64(2): 241-247.
- Gustman, Alan L. and Thomas L. Steinmeier. 1993. "Pension Portability and Labor Mobility: Evidence from the Survey of Income and Program Participation" *Journal of Public Economics*, 50(3): 299-323.
- Hansen, Stephen and Wim Van Der Stede. 2004. "Multiple Facets of Budgeting: An Exploratory Analysis" *Management Accounting Research*, 15(4): 415-439.
- Harris, M. and B. Holmstrom. 1982. "A Theory of Wage Dynamics" *The Review of Economic Studies*, 49(3): 315-333.
- Hinkin, T. R. and J. B. Tracey. 2000. "The Cost of Turnover: Putting a Price on the Learning Curve" *Cornell Hotel and Restaurant Administration Quarterly*, 41(3): 14.
- Hirsch, W. Z. 1956. "Firm Progress Ratios" *Econometrica*, 24(2): 136-143.
- Holmstrom, Bengt. 1982. "Moral Hazard in Teams" *The Bell Journal of Economics*, 13(2): 324-340.
- Jovanovic, Boyan. 1979a. "Job Matching and the Theory of Turnover" *Journal of Political Economy*, 87(5): 972.
- Jovanovic, Boyan. 1979b. "Firm-Specific Capital and Turnover" *Journal of Political Economy*, 87(6): 1246.
- Kahn, C. and G. Huberman. 1988. "Two-Sided Uncertainty and" Up-Or-Out" Contracts" *Journal of Labor Economics*, 6(4): 423-444.
- Kahn, Lisa. 2008. "Asymmetric Information between Employers" Department of Economics, Harvard University.mimeo.Cambridge, MA.
- Katz, Eliakim and Adrian Ziderman. 1990. "Investment in General Training: The Role of Information and Labour Mobility" *The Economic Journal*, 100(403): 1147-1158.

- Krashinsky, Harry. 2002. "Evidence on Adverse Selection and Establishment Size in the Labor Market" *Industrial and Labor Relations Review*, 56(1): 84-96.
- Lazear, E. P. 1979. "Why is there Mandatory Retirement?" *Journal of Political Economy*, 87(6): 1261.
- Lazear, E. P. and S. Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts" *The Journal of Political Economy*, 89(5): 841-864.
- Lazear, Edward P. 1981. "Agency, Earnings Profiles, Productivity, and Hours Restrictions" *The American Economic Review*, 71(4): 606-620.
- Loewenstein, M. A. and J. R. Spletzer. 1998. "Dividing the Costs and Returns to General Training." *Journal of Labor Economics*, 16(1): 142-171.
- Mace, D. and E. Yoder. 2009. *Federal Employees' Almanac: Federal Employees News Digest*.
- Manchester, Colleen F. 2008. "The Effect of Tuition Reimbursement on Turnover: A Case Study Analysis." In *In the Analysis of Firms and Employees: Quantitative and Qualitative Approaches*, ed. Stefan Bender, Julia Lane, Kathryn L. Shaw, Fredrick Andersson, and Till Von Wachter. Chicago: University of Chicago Press.
- Manchester, Colleen F. 2010. "Investment in General Human Capital and Turnover Intention" *American Economic Review*, 100(2): 209.
- McCue, K. 1996. "Promotions and Wage Growth" *Journal of Labor Economics*, 14(2): 175-209.
- Milgrom, P. and S. Oster. 1987. "Job Discrimination, Market Forces, and the Invisibility Hypothesis" *The Quarterly Journal of Economics*, 102(3): 453-476.
- Milgrom, P. and J. Roberts. 1982. "Limit Pricing and Entry Under Incomplete Information: An Equilibrium Analysis" *Econometrica*, 50(2): 443-459.
- Milgrom, Paul. 1988. "Employment Contracts, Influence Activities, and Efficient Organization Design" *The Journal of Political Economy*, 96(1): 42-60.
- Milgrom, Paul and John Roberts. 1988a. "Communication and Inventory as Substitutes in Organizing Production" *The Scandinavian Journal of Economics*, 90(3, Information and Incentives. Vol. 1: Organizations and Markets): 275-289.

- Milgrom, Paul and John Roberts. 1988b. "An Economic Approach to Influence Activities in Organizations" *The American Journal of Sociology*, 94(Supplement: Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure): S154-S179.
- Mincer, J. and B. Jovanovic. 1981. "Labor Mobility and Wages" *Studies in Labor Markets*.
- Montgomery, Edward and Kathryn Shaw. 1997. "Pensions and Wage Premia" *Economic inquiry*, 35(3): 510-522.
- Morita, Hodaka. 2001. "Choice of Technology and Labour Market Consequences: An Explanation of U.S.-Japanese Differences" *The Economic Journal*, 111(468): 29-50.
- Mortensen, D. T. 1978. "Specific Capital and Labor Turnover" *The Bell Journal of Economics*, 9(2): 572-586.
- Munasinghe, L. and B. O'Flaherty. 2005. "Specific Training Sometimes Cuts Wages and always Cuts Turnover" *Journal of Labor Economics*, 23(2).
- Novos, I. E. 1995. "Imperfections in Labor Markets and the Scope of the Firm" *International Journal of Industrial Organization*, 13(3): 387-410.
- Owan, H. 2004. "Promotion, Turnover, Earnings, and Firm-Sponsored Training" *Journal of Labor Economics*, 22(4): 955-978.
- Pinkston, J. 2008. "A Model of Asymmetric Employer Learning with Testable Implications" *Review of Economic Studies*, forthcoming.
- Prendergast, C. 1993. "A Theory of "Yes Men"" *The American Economic Review*: 757-770.
- Prendergast, C. and R. H. Topel. 1996. "Favoritism in Organizations" *Journal of Political Economy*, 104(5).
- Ricchetti, Bryan. 2007. "Turnover as a Gateway to Symmetric Information: Explaining Patterns of Entry in Personnel Data" Unpublished.
- Rosen, Sherwin. 1982. "Authority, Control, and the Distribution of Earnings" *The Bell Journal of Economics*, 13(2): 311-323.

- Salop, Joanne and Steven Salop. 1976. "Self-Selection and Turnover in the Labor Market" *The Quarterly Journal of Economics*, 90(4): 619-627.
- Saporta, I. and M. Farjoun. 2003. "The Relationship between Actual Promotion and Turnover among Professional and Managerial-Administrative Occupational Groups" *Work and Occupations*, 30(3): 255.
- Sattinger, M. 1975. "Comparative Advantage and the Distributions of Earnings and Abilities" *Econometrica*, 43(3): 455-468.
- Schonberg, Uta. 2007. "Testing for Asymmetric Employer Learning" *Journal of Labor Economics*, 25(4): 651-691.
- Shaw, Kathryn and Edward P. Lazear. 2007. "Tenure and Output" *Labour Economics*, In Press, Corrected Proof.
- Song, Y. 2007. "Recall Bias in the Displaced Workers Survey: Are Layoffs really Lemons?" *Labour Economics*, 14(3): 335-345.
- Stevens, Ann H. 1997. "Persistent Effects of Job Displacement: The Importance of Multiple Job Losses" *Journal of Labor Economics*, 15(1, Part 1): 165-188.
- Topel, R. H. and M. P. Ward. 1992. "Job Mobility and the Careers of Young Men" *The Quarterly Journal of Economics*, 107(2): 439-479.
- Waldman, Michael. 1984a. "Job Assignments, Signalling, and Efficiency" *The Rand journal of economics*, 15(2): 255-267.
- Waldman, Michael. 1984b. "Worker Allocation, Hierarchies and the Wage Distribution" *The Review of Economic Studies*, 51(1): 95-109.
- Waldman, Michael. 2008. "Theory and Evidence in Internal Labor Markets." In *Handbook of Organizational Economics.*, Forthcoming ed., ed. R. Gibbons and J. Roberts.
- Wasmuth, W. J. and S. W. Davis. 1983. "Strategies for Managing Employee Turnover" *Cornell Hotel and Restaurant Administration Quarterly*, 24(2): 65.
- Yellen, J. L. 1984. "Efficiency Wage Models of Unemployment" *The American Economic Review*, 74(2): 200-205.

Zhang, Ye. 2007. "Employer Learning Under Asymmetric Information: The Role of Job Mobility".