SIEPR Discussion Paper No. 02-16

Labor Market Competition using Compensation Schemes and Intertemporal Relationships

By
Jeremy T. Fox
Stanford University
November 2002

Stanford Institute for Economic Policy Research
Stanford University
Stanford, CA 94305
(650) 725-1874

The Stanford Institute for Economic Policy Research at Stanford University supports research bearing on economic and public policy issues. The SIEPR Discussion Paper Series reports on research and policy analysis conducted by researchers affiliated with the Institute. Working papers in this series reflect the views of the authors and not necessarily those of the Stanford Institute for Economic Policy Research or Stanford University.
Labor Market Competition
using
Compensation Schemes and Intertemporal Relationships

Jeremy T. Fox∗

November 21, 2002

Abstract
It is commonly thought that firms use aspects of compensation schemes, such as wage levels and promises of future raises, to reduce worker turnover, attempting to increase productivity through intertemporal relationships (such as human capital investments and dynamic incentive schemes) with their workers. Using data on Swedish white collar workers, I estimate a dynamic labor supply model to measure the empirical relevance of the belief that compensation schemes have large impacts on worker turnover. The model allows each worker to choose between staying at his current firm and moving to a particular new one, according to the compensation schemes offered by all firms. Workers consider the expected value of the wage profiles offered by each firm in addition to moving costs. The results show that the sensitivity of worker turnover to compensation schemes is low and it is not cost-effective for firms to reduce turnover by increasing future raises. Older workers rarely change employers because they face high movement costs and do not receive large wage increases upon moving. However, the turnover behavior of younger workers is more sensitive to future raises and the wages of young workers change more upon moving. These findings are consistent with a story where firms compete for younger workers much more than for older workers.

∗Thanks go to my advisers and the other faculty at Stanford. This paper is part of the project Pay and Promotion in Swedish Private Industries 1970-1990, which is under the leadership of Eva Meyerson-Milgrom, with financial support from the Swedish Council for Research in the Humanities and Social Sciences. I am personally supported by The Kapnick Foundation through a grant to the Stanford Institute for Economic Policy Research. My email address is jerfox@stanford.edu. This is my job market paper.
## Contents

1 Introduction ................................. 1

2 Empirical Evidence on Potential Labor Market Competition and Worker Turnover 3

3 Theory of Labor Market Competition and Turnover 6
   3.1 Worker’s Problem (Labor Supply) .............................. 6
   3.2 Firm’s Problem (Labor Demand) .............................. 9

4 Labor Supply Estimation Procedure 12
   4.1 Estimation Overview .............................. 13
   4.2 Discussion of the Labor Supply Model .............................. 13
   4.3 Compensation Schemes and States .............................. 15
   4.4 Overall Identification Strategy .............................. 17

5 Data .............................. 21
   5.1 Overview .............................. 21
   5.2 Firm Estimation Sample .............................. 22
   5.3 Worker Estimation Sample .............................. 23
   5.4 Worker Turnover Across Estimation Groups .............................. 23

6 Compensation Schemes .............................. 26
   6.1 Compensation Schemes for Firm Stayers .............................. 28
   6.2 Compensation Schemes for Firm Movers .............................. 32

7 Results for Labor Supply .............................. 36
   7.1 Parameter Estimates .............................. 36
   7.2 Model Fit .............................. 43
   7.3 Turnover Marginal Effects .............................. 45
   7.4 Robustness: Firm-Specific Movement Costs .............................. 49

8 Discussion of the Results .............................. 50
   8.1 Overview of the Empirical Results .............................. 50
   8.2 Informed Speculation About Firm Behavior in Labor Markets .............................. 51
   8.3 An Explanation for Concave Age-Earnings Profiles? .............................. 53

9 Conclusions .............................. 53
1 Introduction

The burgeoning fields of organizational and personnel economics demonstrate how firms use complex policy instruments, such as compensation schemes, to increase the performance of their workers. For example, firms may motivate workers to put forth more effort by giving future raises which are conditional on hard work now. (Becker and Stigler, 1974; Lazear, 1979) It is not always clear how these models relate to the common view that competition in the labor market for workers is a strong force that eliminates wage differentials across firms. For example, a firm trying to recruit workers may need to offer high current period wages and may also thus find it hard to afford large future raises for an incentive device. Additionally, workers may simply avoid a firm with an especially intense work environment. In these cases I say that labor market competition constrains the firm’s choice of compensation schemes and the level of worker productivity achieved through its intertemporal incentive scheme. In this paper, I address whether the need to recruit and retain labor prevents a firm from implementing the first best internal arrangements envisioned in the organizational literature. My investigation presents complicated measurement issues, which I address by estimating a dynamic model of labor supply to particular firms.

The primary question I ask is how the level of competition in the external labor market affects the level of worker productivity achieved by intertemporal employment relationships, such as human capital investment or dynamic incentive schemes. These types of relationships involve investments of time and effort on the part of both the worker and firm that are, to some degree, specific to the employment match. Because I do not observe worker productivity, I use worker turnover, which disrupts intertemporal relationships, as a proxy for the potential effectiveness of an employment match. Investments in employment relationships in a high-turnover firm are probably not as profitable as investments in low-turnover firms, where individual employment relationships are more likely to last. In the absence of compensation changes, intuition suggests that increases in external labor market competition - for instance, the entry of new firms - will induce additional worker turnover. However, because economists do not know whether or how a firm adjusts its compensation scheme in reaction to external competition, we are unable to predict the effects of such competition on worker turnover. Firms may react to increased external competition by either deemphasizing the use of intertemporal relationships or by highlighting such relationships through modified compensation schemes designed to reduce turnover. In this paper, I estimate whether firms can cost-effectively reduce turnover through compensation schemes by estimating the sensitivity of worker turnover to increases in future raises with a dynamic model of labor supply to individual firms.

Using a rich data source that includes information on the career profiles of individual workers and assuming that workers are non-strategic actors, I consistently estimate labor supply curves facing individual firms. With these estimates, I explicitly calculate the sensitivity of worker turnover and other worker behavior to the compensation schemes of all firms in the labor market. These results determine
whether it is cost-effective for a firm to reduce turnover to its particular desired level or too expensive, meaning that turnover is largely exogenous from a firm’s point of view. I also use my estimates to test hypotheses about the variations in worker sensitivity to compensation schemes according to worker age. Lastly, I use my labor supply estimates and descriptive data on firms’ compensation schemes to put forward suggestions regarding the motives and strategies used by firms competing in the labor market.

Several theories suggest that the slope of the age-wage profile is a tool that can be used to both reduce turnover and improve worker productivity through raises that are conditional on exerting effort. (Salop and Salop, 1976; Becker and Stigler, 1974; Lazear, 1979) As a consequence, the labor supply model that I specify and estimate by maximum likelihood is a discrete choice, dynamic programming model of workplace selection. Workers consider the age-wage profiles and nonwage amenities offered by each firm. They also face moving costs which dissuade them from always choosing the employer offering the highest current period wages. Age-specific switching costs and finite career horizons are two reasons why older workers may behave differently from younger workers. Even though workers switch firms on the equilibrium path, the dynamic programming model properly weights each potential future outcome by the probability that it might happen.

My analysis of the effects of labor market competition on compensation schemes and intertemporal relationships requires comprehensive data. In order to estimate the observed compensation schemes of individual firms, I need data describing multiple workers from the same firms competing in one labor market. Likewise, in order to estimate my firm choice model, I need data that tracks workers as they switch employers. Finally and perhaps most importantly, to consistently estimate a labor supply function, I need data on employers not chosen by particular workers, or in other words the outside options. These three data requirements are met by data compiled by the Swedish Employers’ Federation (SAF). Over a 21 year period, the SAF collected information on the wages and other characteristics of a large fraction of private sector workers in Sweden.1 Because I am interested in intertemporal relationships, I focus only on white collar workers, for whom models of long-term employment relationships are most relevant. Furthermore, because I am most interested in describing the effects of labor market competition on compensation schemes, I want to focus on labor markets that are as homogeneous in observable worker characteristics as possible. For this reason, I consider separately several labor markets for full-time, male workers with particular educational backgrounds.

In the rest of the paper, I first motivate the importance of studying labor market competition and worker turnover using descriptive empirical work and a theoretical model. I then estimate a structural labor supply model and interpret the results.

My main conclusion is that the sensitivity of worker turnover to compensation schemes is low and

---

1Despite Sweden’s reputation as a country where national wage bargaining has reduced wage heterogeneity, Swedish firms do use different wage policies. In 1990, the firm at the 75th percentile paid 10% more to 25 year old, male, full-time, engineers than the firm at the 25th percentile in the wage distribution. Wage policies also vary the rate at which wages increase with time at a firm. For the five year wage increases that engineers aged 25 to 29 in 1985 received between 1985 and 1990, the 90th percentile firm on this metric gave a 1.7% higher nominal wage increase than the 10th percentile firm.
it is not cost-effective for firms to reduce turnover by increasing wages. This is especially true for older workers, who face high movement costs and short remaining careers. Job changing by older workers is thus largely driven by idiosyncratic reasons and not compensation schemes. Younger workers are more likely to change employers because they have lower movement costs and longer remaining careers to enjoy wage gains. This evidence about worker behavior coincides well with additional evidence about the equilibrium compensation scheme choices by firms. A young firm mover tends to have a larger fractions of his new wage based upon the compensation scheme of his new firm and a smaller fraction based upon his experience at his previous employer. Also, the compensation schemes offered to younger, but not older, workers tend to be correlated with empirical measures of potential labor market competition. Altogether, my findings about worker and firm behavior are consistent with a story where firms compete aggressively for younger workers who then become locked into careers at one firm.

2 Empirical Evidence on Potential Labor Market Competition and Worker Turnover

In this section I provide motivating evidence on the descriptive empirical relationship between potential labor market competition and worker turnover. In Sweden, several measures of potential local labor market competition are correlated with higher worker turnover. Due to their relative simplicity, and perhaps greater degree of exogeneity from the point of view of individual workers and firms, I use the population densities of Swedish counties as my measure of potential labor market concentration. In addition to demonstrating the well-known empirical fact that turnover declines with worker age (Farber, 1999), Table 1 shows that worker turnover for two different educational groups is higher in more densely populated Swedish counties for most age groups. For example, engineers aged 40-44 have a 0.027 higher probability of switching employers in a densely populated county. The main exception is young engineers, who change firms at a high rate, but do not appear to behave differently across county types.\footnote{The positive correlation between population density and worker turnover extends to other measures of potential labor market competition, such as the number of firms competing for labor in a given region and the Herfindahl index of employment share concentration. These descriptive correlations were removed from the paper for space reasons. They are available from me upon request.}

One concern with these averages is that they do not control for widely known empirical patterns involving turnover. In particular, larger firms have lower turnover and certain industries, such as construction, have higher turnover than others. In Table 2, I report results from a linear probability regression of whether a worker moves over a five year period on the log of his county’s population density interacted with worker age. I also include the unreported controls log of a worker’s establishment’s size and indicator variables for worker ages, calendar years and establishment industries. It is clear that there is a statistically significant correlation between turnover and population density, even after
Table 1: Five Year Turnover Probabilities for Swedish White Collar Workers by County Population Density and Starting Worker Age and Education\textsuperscript{a}

<table>
<thead>
<tr>
<th>Starting Age Group</th>
<th>Vocational High School</th>
<th>University Engineers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Densely Pop. Counties</td>
<td>Sparsely Pop. Counties</td>
</tr>
<tr>
<td>25-29</td>
<td>0.123</td>
<td>0.096</td>
</tr>
<tr>
<td>30-34</td>
<td>0.105</td>
<td>0.076</td>
</tr>
<tr>
<td>34-39</td>
<td>0.081</td>
<td>0.061</td>
</tr>
<tr>
<td>40-44</td>
<td>0.062</td>
<td>0.041</td>
</tr>
<tr>
<td>45-49</td>
<td>0.042</td>
<td>0.028</td>
</tr>
<tr>
<td>50-54</td>
<td>0.039</td>
<td>0.016</td>
</tr>
<tr>
<td>55-59</td>
<td>0.028</td>
<td>0.014</td>
</tr>
<tr>
<td>Total N</td>
<td>154,243</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a} These means are over individual workers employed by firms in a panel that I describe in Section 5. The 3 densely populated counties are Stockholm, Gothenburg and Malmo. These are also the counties with the largest populations, by far. There are 21 sparsely populated counties. The averages are taken over five year intervals beginning in 1970 and ending in 1990. Only male white collar workers who work more than 35 hours a week are included in this table. The overall (ignoring age) differences in turnover between county types are statistically significant at any reasonable level.
Table 2: Linear Probability Models for Moving Over a five year Period on Population Density Interacted with Age while Controlling for Other Known Factors Correlated with Worker Turnover

<table>
<thead>
<tr>
<th>Starting Age Group</th>
<th>Vocational</th>
<th>University</th>
<th>Engineers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Log of Pop. Density</td>
<td>Standard Error</td>
<td>Log of Pop. Density</td>
</tr>
<tr>
<td>25-29</td>
<td>0.0207</td>
<td>0.0025</td>
<td>0.0034</td>
</tr>
<tr>
<td>30-34</td>
<td>0.0196</td>
<td>0.0022</td>
<td>0.0105</td>
</tr>
<tr>
<td>34-39</td>
<td>0.0149</td>
<td>0.0020</td>
<td>0.0165</td>
</tr>
<tr>
<td>40-44</td>
<td>0.0132</td>
<td>0.0020</td>
<td>0.0179</td>
</tr>
<tr>
<td>45-49</td>
<td>0.0103</td>
<td>0.0019</td>
<td>0.0122</td>
</tr>
<tr>
<td>50-54</td>
<td>0.0146</td>
<td>0.0020</td>
<td>0.0155</td>
</tr>
<tr>
<td>55-59</td>
<td>0.0089</td>
<td>0.0025</td>
<td>0.0134</td>
</tr>
<tr>
<td>Total N</td>
<td>154,243</td>
<td></td>
<td>107,686</td>
</tr>
</tbody>
</table>

Moving is defined as switching establishments within a five year period. Other variables included are the log of initial establishment size and indicator variables for industry, worker age, and calendar year. The population density information is for 1990. The sample is identical to the one in Table 1. The standard errors are robust to heteroskedasticity, and use clustering at the individual worker level.

holding the other controls constant. For example, the coefficient on population density of 0.0165 for engineers with a starting age of 34-39 means that a county with twice the population density of another has a statistically predicted five year turnover rate that is 0.011 higher. The average turnover rate for engineers for this sample is 0.177.

These descriptive results are consistent with a theory that the number of outside options available to a worker increases the chance that he will switch employers in a given period. Some readers may find this positive correlation between measures of potential labor market competition and worker turnover unsurprising. Others familiar with the theoretical literature on how firms can reduce turnover by altering compensation schemes may find this result more provocative. It appears firms in densely populated Swedish counties choose not to reduce turnover to the level found in sparsely populated counties either because it is not cost-effective in terms of wages or because the worker productivity benefits from reducing turnover on the margin are low. In both cases, this empirical relationship supports the need to consider labor market competition when studying worker turnover.

---

3Sweden’s population as a whole is relatively concentrated in the southern part of the country, so I do not want to push the assumption that Swedish counties are distinct local labor markets too far.
3 Theory of Labor Market Competition and Turnover

This section outlines how workers make labor supply decisions and how firms balance competing objectives in internal and external labor markets. All other things held equal, workers may be more likely to change jobs when competition in the external labor market increases. Turnover decreases the ability of intertemporal relationships to raise the productivity level of workers. The key question then becomes whether firms can cost-effectively adapt their compensation schemes to reduce turnover enough to maintain their intertemporal relationships. The effects of compensation scheme changes may vary with the age of a worker. I present a simple model to illuminate these main points.

The theoretical labor supply model described in this section is the backbone of the structural model that I estimate in later sections. I present the labor supply model here first in order to highlight the economic issues without confounding the discussion with details of empirical implementation.

3.1 Worker’s Problem (Labor Supply)

In the model, the main decision of a worker is to choose an employer \( j \) out of a total number of firms \( J \). Workers are employed until age 65, which is the mandatory retirement date in the data. The intermediate ages are indexed by \( t \). Let worker \( i \) of age \( t \) and state variable \( s_{it} \) have a current period utility from accepting a job at firm \( j_{it} \) of

\[
u(s_{it}, j_{it}) + \epsilon_{ijt} = \beta_w w_{ijt} + x(j_{i,t-1}, j_{it})' \beta_{xt} + \xi_j + \epsilon_{ijt}.
\]

(1)

Utility is equal to the sum of the (log) wage \( w_{ijt} \) times a parameter \( \beta_w \); a vector of other observable attributes \( x(j_{i,t-1}, j_{it}) \) times an age-specific parameter vector \( \beta_{xt} \), which together address various measures of moving costs between the old firm \( j_{i,t-1} \) and the new firm \( j_{it} \); time and worker invariant nonwage attributes of the job \( \xi_j \); and person-job specific preference shocks \( \epsilon_{ijt} \). The movement costs \( x(j_{i,t-1}, j_{it})' \beta_{xt} \) are important, because without movement costs a worker will instantaneously switch from firm to firm in order to obtain the highest utility benefit each period. With movement costs, each worker must carefully consider the tradeoff between incurring movement costs and receiving future wage increases. Movement costs depend on age because a young worker just leaving his parents’ house or his college dormitory probably has a much easier time of moving than an older worker with dependents who has established long-term ties in a community and workplace. One way of decomposing the nonwage attributes is \( \xi_j = d_j - e_j \), where \( d_j \) captures characteristics of the job, such as whether the product being made is especially interesting or challenging, and \( e_j \) is the work effort level, which enters utility negatively. Firm heterogeneity in \( d_j \) may cause compensating wage differentials, while some of the moral hazard ideas in the literature on intertemporal relationships are captured by \( e_j \).\(^4\)

\(^4\)In a more general model, \( \xi_j \) could vary by worker age as well. I choose not to incorporate time-varying fixed effects in the estimation of the model for computational reasons.
A worker picks the employer that maximizes his expected discounted future utility

\[
E\left[ \sum_{\tau=t}^{65} \delta^{\tau-t} \left( u(s_{i\tau}, j_{i\tau}) + \epsilon_{i\tau} \right) \mid s_{it}, \epsilon_{it} \right],
\]

where \( \delta \) is a discount factor between 0 and 1, and \( \epsilon_{i\tau} \) is the vector of all firm-specific shocks \( \epsilon_{i\tau} \). Given that a worker has a finite career-horizon, there will be less time to earn positive utility increases from a job move as the worker ages.

The state variable \( s_{it} \) in this case is simply a vector composed of the worker’s age \( t_i \), the employer \( j_{it-1} \) where the worker was employed last period, and last period’s wage \( w_{i,t-1,t-1} \). The wage of a worker relative to peers of his age at his old employer is my measure of the worth of a worker relative to his co-workers, and is a sufficient statistic for his past labor market outcomes. The state vector \( s_{it} \) follows the transition rule \( f(s_{i,t+1} \mid s_{it}, j_{it}, \theta_p) \). Because age, \( t_i \), is deterministic and the employer, \( j_{it} \), is a choice of the worker,

\[
f(s_{i,t+1} \mid s_{it}, j_{it}, \theta_p) = f(w_{it} \mid w_{i,t-1}, j_{it}, j_{it-1}, t_i),
\]

the transition rule for wages alone. The major intertemporal uncertainty in a worker’s decision problem is how the random part of the state variable, the wage, changes from year to year. I call a firm’s wage transition probabilities its compensation scheme.\(^5\) As an example of how this formulation for compensation schemes works, firms that put greater probabilities on higher raises \((w_{ijt} - w_{ij,t-1})\) have steeply sloped age-wage profiles. Also, firms can treat workers moving from other firms differently than workers who stay at the current firm.\(^6\)

The worker’s problem can be solved by the dynamic programming technique of backwards recur-

---

\(^5\)Based upon the notation I have chosen, it should be clear that wages today are a function only of wages last period and current period choices made by the worker. This means that the stochastic transition rule for wages is Markovian, requiring that wages depend directly on only state variables last period. This does not rule out the possibility that workers with low seniority at a particular job (relative to their age) are paid differently than workers with high seniority, if the starting wages of older workers who join a new firm are lower or higher than workers of the same age with more seniority. On the other hand, the Markovian wage process does rule out different future wage distributions for two workers who previously were paid the same wage.

\(^6\)This compact representation \( f(w_{it} \mid w_{i,t-1}, j_{it}, j_{i,t-1}, t_i) \) for the compensation scheme contains quite a bit of economics. Unfortunately, I will not be able to achieve all of this flexibility with the particular empirical specifications I employ, yet readers interested in applying this type of technique themselves should be interested in knowing the generality of the framework. An age-wage profile with a steep slope is captured by placing large amounts of probability mass on \( w_i \)'s that are much higher than the current \( w_{ij,t-1} \). The slope can vary by the stage of the lifecycle, as age-wage profiles tend to have a concave shape. A firm where high values of \( w_{ij,t-1} \) correlate with high values for \( w_{ij} \) is operating a fast track, where the firm separates workers into winners and losers. Other firms might use a regression-to-the-mean policy, where those earning too much for their age group are given lower raises, while those falling behind are given more. Additionally, a firm can treat workers considering joining the firm differently than current employees and independently of each other depending on the identity of their former employers. If certain aspects of employment relationships are specific to the current employer, then firms may discriminate against incoming workers. On the other hand, new workers could receive higher wages as part of a competitive response in order to recruit labor from a particular labor market rival. Another possibility is that new workers from a particular firm could receive less as part of an understanding that the old employer attracts mediocre workers, and that it is unprofitable to pay those workers high wages.
Define the continuation value to be
\[
V(s_{it}) = \int \max_{j_{it} \in \{1, \ldots, J\}} \mathbb{E} \left[ \sum_{\tau=t}^{65} \delta^{\tau-t} (u(s_{i\tau}, j_{i\tau}) + \epsilon_{i\tau}) \mid s_{it}, \epsilon_{it} \right] g(\epsilon_{it})
\]
\[
= \int \max_{j_{it} \in \{1, \ldots, J\}} \left[ u(s_{it}, j_{it}) + \epsilon_{iit} + \delta \sum_{s_{i,t+1} \in S} V(s_{i,t+1}) p(s_{i,t+1} \mid s_{it}, j_{it}, \theta_p) \right] g(\epsilon_{it})
\]
\[
= \int \max_{j_{it} \in \{1, \ldots, J\}} [v(s_{it}, j_{it}) + \epsilon_{iit}] g(\epsilon_{it}),
\]
where \(g\) is the joint density of the shocks \(\epsilon_{it}\), and where \(v(s_{it}, j_{it})\) are defined to be choice-specific value functions. If the worker-specific shock has the logit distribution, then the probability of picking firm \(j_{it}\) for an aged \(t\) worker is
\[
P(j_{it} \mid s_{it}) = \frac{\exp(v(s_{it}, j_{it}))}{\sum_{k=1}^I \exp(v(s_{it}, k_{it}))}.
\]
The higher the utility \(v(s_{it}, j_{it})\) from any given choice, the more likely it is to be picked. Also, the more attractive other firms are, the less likely a particular firm is to be picked: the denominator of the choice probabilities increases if the competitor firms become more attractive employers. For this paper, an important but obvious result is that if the number of other firms in the labor market increase, the probability of picking any one firm will decrease, if factors chosen by firms, such as compensation schemes, are held constant. This happens because the sum of the utilities from all other employers enter the denominator of the choice probability for each firm.

In this framework, all firms will suffer turnover with positive probability because of the presence of the logit error terms. However, firms that are more attractive to their current workers will have less turnover, and those that are attractive to workers at other firms will recruit workers with higher probability. If there are \(I\) total workers in the labor market, the sum
\[
\sum_{i=1}^I P(j_{it} \mid s_{it})
\]
is one measure of the expected number of workers hired by firm \(j\). If there are \(I_{jt}\) workers at firm \(j\) of age \(t\) last period, the sum
\[
\frac{1}{I_{jt}} \sum_{i=1}^{I_{jt}} (1 - P(j_{it} \mid s_{it}))
\]
represents the average probability of a worker at the firm turning over in any given year. For an individual worker, this is just 1 minus the probability of staying at the firm. For the turnover statistic, the sum is calculated over only those workers of age \(t\) who were previously employed by firm \(j\). Both recruitment and turnover are properties of the more abstract relation that I call the labor supply curve.

In this dynamic model, the decision of which firm to work at today is influenced by the time path
of future wages. Firms must compete in both the current spot market and in the futures market, where they offer distributions of the future wages that may arise.

The model that I present here seeks to be consistent with certain explanations that economists offer for why observationally equivalent workers are paid different wages in different firms. The three main explanations are that there is some unmeasured heterogeneity in worker abilities, that workers receive compensating differentials for unpleasant work environments, and that some firms pay efficiency wages in order to increase work effort or reduce turnover. (Rosen, 1974, 1986; Shapiro and Stiglitz, 1984; Murphy and Topel, 1987) As a way of addressing worker heterogeneity, I let the past wage of a worker be part of his state variable and thus a determinant of the distribution of future wage outcomes. I do not attempt to explicitly measure a distribution of unobserved worker abilities.7 From the perspectives of a paper where estimating labor supply curves is a major goal, to the extent that a firm can adjust the level of its nonwage attractiveness to workers and the amount of costly effort it asks its workers to perform, compensating wage differentials and efficiency wages are two sides of the same coin. In my model, effort is measured by $e_j$ and the nonwage benefits of a firm are captured by $d_j$. Both are subsumed into a fixed effect $\xi_j = d_j - e_j$, because both effort and nonwage attractiveness are factors that enter worker utility and are unobserved to the econometrician.

3.2 Firm’s Problem (Labor Demand)

My goal in this paper is to use empirical analysis to determine how firms balance the multiple objectives of recruiting, retaining, and motivating labor. Unfortunately, I do not observe a measure of the use of productivity-enhancing intertemporal relationships. Thus I restrict the estimation to produce only the sensitivity of worker turnover to changes in a firm’s compensation scheme. While I do not estimate the parameters entering into the firm’s problem because I am unsure of the exact maximization problem facing firms in the unionized Swedish economy, it is useful to briefly discuss how firms make their decisions in order to understand why studying worker turnover is important.

Firms compete in the labor market by offering a package consisting of a compensation scheme $f(w_{it} | w_{i,t-1}, j_{it}, j_{i,t-1}, t_i)$ and a working environment, captured by the fixed effect $\xi_j = d_j - e_j$. The firm chooses these variables to maximize expected profits

$$E[\pi_j] = \sum_{\tau=0}^{\infty} \delta^\tau E \left[ \sum_{i=1}^{I} P(j_{i\tau} | s_{i\tau}) \left( y(e_j, \omega_j) - \exp(w_{ij\tau}) \right) \right],$$

which is just the discounted future value of the difference between a worker’s output and his cost times the probability of being able to hire the worker. Here $y(e_j, \omega_j)$ is the output of a worker exerting effort

---

7My earlier descriptive work on age-wage profiles in Sweden suggests that fixed and known abilities are not a good explanation for the patterns of wages for observationally equivalent workers, although learning about ability could be important. Fox (2002)
and $\omega_j$ is a technology parameter that reflects how sensitive worker productivity is to effort. Finally, $w_{ijt}$ is the log of the wages paid to such a worker.

In Sweden it is very difficult to fire a worker or to reduce his wages, so it is likely that the main way firms encourage their workers to exert effort is to promise them raises and promotions in the future. This type of dynamic incentive relationship can be conceptualized by requiring that all compensation schemes and effort requests be incentive compatible. Incentive compatibility means that the benefits to the worker from exerting the requested effort $e_j$ exceed the punishment he receives from shirking. In personnel economics and efficiency wage models the punishment is often arbitrary; in this case it is not receiving a raise. Thus, for a given worker, the incentive compatibility constraint is

$$-\frac{e_j}{\delta} + \sum_{s_{t+1} \in S} V(s_{t+1}) f(s_{it,t+1} \mid s_{it}, j_{it}, \theta_P, e \geq e_j) \geq \sum_{s_{t+1} \in S} V(s_{t+1}) f(s_{it,t+1} \mid s_{it}, j_{it}, \theta_P, e < e_j). \tag{6}$$

This constraint has been calculated by listing on the left hand side the disutility from effort and the benefits of future raises from working as the firm requests, while the right hand side shows what happens when a worker is refused raises for not putting forth effort. The wages from the current period’s job are not in themselves important, because the worker receives them either way. This is another way of restating my earlier point that high current wages by themselves cannot induce effort in an economy where firing is difficult. As a further point, in a low-turnover setting the present discounted value of utility will also be higher for younger workers. This is evidence that younger workers may be more receptive to productivity-enhancing dynamic incentive schemes.

As equation 4 shows, when the number of firms in the market increase, the probability of picking any one firm decreases, if all other factors are held equal. This shows that the probability of changing employers, equation 5, rises when the number of firms increases. If turnover probabilities are very high, then the incentive compatibility constraint, equation 6, is less likely to hold for a given value of effort, as the effect of any punishment for shirking is dependent on the worker remaining at the firm. The implication is that firms in competitive labor markets with high turnover have a hard time implementing intertemporal relationships. High turnover gives a firm an incentive to adjust its compensation schemes in order to reduce turnover.

---

8More generally, $\omega_j$ could capture other aspects of a firm’s production function that are related to the need to reduce turnover, such as the sensitivity of output to the level of specific training held by workers. Some firms with highly distinctive production technologies may have increased levels of $\omega_j$. It is not critical to specify what exactly this variable represents, as long as it doesn’t enter in the utility function of workers in a manner not captured by the fixed effect $\xi_j$.

9Remember, the term $t$ refers to the age of a worker. At any given firm, there will be workers of many different ages.

10This might be the case if firms are restricted to offering straight monthly salaries that cannot be decreased, as is the case in Sweden for many workers. More formally, this type of raise and promotion scheme (or some bonus scheme) is optimal whenever worker output (and hence shirking) cannot be observed until after current-period wages are paid. My timing assumption about the observability of effort rules out piece-rates or current-period bonuses as mechanisms to correct the moral hazard problem.

11Clearly it is unlikely that one $e_j$ will satisfy the incentive compatibility constraint for every single worker in the labor market. I do not let $e_j$ vary within a firm for computational reasons in the estimation procedure.
The multiple goals of a firm should be clear. The firm can recruit more employees by paying higher wages or by promising high future raises. These mechanisms also reduce turnover. Firms often seek to increase effort from their workers by offering large raises, but risk driving them away if the hard work is not adequately compensated with high wages. Finally, despite all these benefits from high wages, firms need to keep their wage bill low in order to earn profits.

Thinking about extreme market conditions can illuminate how labor market competition affects the use of intertemporal relationships. A firm which is a monopsonist in the labor market only needs to pay workers starting wages high enough to motivate them to work. The monopsonist can then create a hierarchy of positions and wage levels such that workers receive raises each period equal to the wage-equivalent cost of the first best level of effort $e^*$, meaning that the firm sets $w_t - w_{t-1} = e^* \delta$ for every age $t$. The monopsonist perfectly addresses its recruitment and retention problems, and can specialize its compensation scheme towards motivating workers to work at the optimal effort level. Conversely, when there is an infinite number of firms in the labor market, workers switch firms with probability 1 each period, regardless of the choices of compensation schemes in the market. All workers change employers because the support of the choice-specific i.i.d. logit error terms is unbounded, and with an infinite number of such terms the deterministic part of worker utility is irrelevant. With an infinite number of firms, employers cannot reduce turnover and cannot motivate workers through the use of intertemporal relationships, because the incentive compatibility constraint does not hold for positive effort levels.

Several previous authors have emphasized steeply sloped age-wage profiles as incentive devices. Salop and Salop (1976) show how age-wage profiles can reduce turnover, while Becker and Stigler (1974) and Lazear (1979) demonstrate that age-wage profiles can improve the effort levels of workers. The latter models that emphasize worker productivity take on more importance in Sweden, because employee-protection policies prevent the layoffs or demotions that are used as punishments in efficiency wage models. Layoffs are legally obliged to occur in the inverse order of job seniority at an employer, and unless a worker is seriously misbehaving he cannot be fired for cause. In this situation, it is the threat of not receiving a raise that motivates workers to put forth more effort, rather than the traditional efficiency wage story in Shapiro and Stiglitz (1984). However, static efficiency wages can still reduce voluntary turnover, as in Stiglitz (1974).

My focus on dynamic models of moral hazard should not preclude other models about screening workers into optimal job assignments or encouraging workers to acquire specific or general human capital. (Becker, 1962) I view all of these models as special cases of intertemporal relationships, by which I mean dynamic, productivity-enhancing arrangements where workers trade some form of effort, whether it be work effort or training effort, in return for some expected future gain. The intertemporal relationship explains why studying worker turnover is important, because changing employers prevents a worker from receiving his reward and thus makes it less likely that he will put forth effort to
Because careers have finite horizons, it is unlikely that labor supply and labor demand decisions are made without considering workers’ ages. Younger workers have a long career ahead of them, and have a large upside to finding the right employer before settling down. Older workers have little time remaining to enjoy higher wages from a move, and may not switch to higher-paying employers if movement costs are positive. Firms have related agendas. Younger workers are malleable material for whom current human capital investments or incentive schemes may provide a large return in increased worker productivity. On the other hand, older workers have a lot of public information about their abilities, have short remaining careers which makes them unlikely to respond to new dynamic incentive schemes and are relatively poor candidates for expensive training programs.

4 Labor Supply Estimation Procedure

In this section I outline an empirical model of how worker turnover is affected by the choice of compensation scheme by a firm. More specifically, I describe details necessary for estimation of the labor supply model in Section 3.1. The quantities that I am interested in calculating are the turnover marginal effects with respect to changes in a compensation scheme. I estimate a structural model because compensation schemes are usually chosen optimally by firms, and simply looking at the empirical correlation between compensation schemes and turnover is not going to provide a reliable estimate of the turnover marginal effects.

In order to see whether firms can entrench intertemporal relationships against outside labor market competition, I need to estimate the sensitivity of worker turnover to changes in the compensation schemes of firms. There are two possible outcomes. If I estimate the effects on turnover from changing compensation schemes to be high, then firms can cost-effectively reduce turnover. Any observed failure to reduce turnover means that the marginal benefits from increased worker productivity through turnover reductions are low. The other possible outcome is that the marginal turnover effects are small. If substantial reductions in turnover involve a very high wage cost, then it might not be profit maximizing to reduce turnover, even if the productivity benefits from intertemporal relationships are not small.

In this case, potential labor market competition constrains the effectiveness of intertemporal relation-

---

12A tradeoff in my particular model is that asking workers to put forth higher effort makes a job itself less desirable and reduces recruitment and increases turnover. The compensation scheme is used to address multiple goals, not just reducing worker turnover. The idea of a tradeoff in compensation policies was explicitly modeled in an earlier paper of mine. (Fox, 2001) However, many papers have presented variations on this type of idea. Jovanovic (1979) shows how expected future turnover will reduce a worker’s investment in specific human capital, although in his model there is no role for the firm to modify worker behavior through its compensation scheme. The literature on training and turnover tends to focus on the idea that all costly investments are made by the employer. In this case, the only role of the compensation scheme is to recruit workers and reduce turnover, without considering the tradeoff between multiple firm objectives. In fact, Waldman (1984), in a world with exogenous levels of specific human capital, demonstrates how training can indirectly reduce turnover by giving a firm a larger budget constraint to ward off competitors.
4.1 Estimation Overview

There are two main stochastic processes in a dynamic discrete choice model: the choice decisions of the agent and the transition probabilities of the state variables. In my model the decision of a worker is a firm \( j \) where to work. My model predicts a choice \( j_{it} \) with probability \( P(j_{it} | s_{it}, \beta) \), which differs based upon the state variables of a worker, \( s_{it} \). The state variables are the age of the worker, the firm the worker chose last period, and the wage the worker received last period. Conditional on a worker’s choice of a firm, the only stochastic state variable is the wage. Thus the transition rule \( f(s_{i,t+1} | s_{it}, j_{it}, \theta_w) \) is equivalent to the compensation scheme of the chosen firm, \( j_{it} \), or \( f(w_{ijt} | w_{i,j_{it-1},t-1}, j_{it}, j_{i,t-1}, t_i) \).

The combination of the decision of the agent and the evolution of the state variable forms the likelihood function

\[
L(\beta, \theta_w | s, j) = \prod_{i=1}^{I} P(j_{it} | s_{it}, \beta) f(s_{i,t+1} | s_{it}, j_{it}, \theta_w),
\]

with observations on \( I \) workers indexed by \( i \). Maximizing the likelihood function gives consistent and efficient estimates for the parameters that enter the utility function of the worker, \( \beta \), and the parameters that enter the state transition probabilities, \( \theta_w \). The only stochastic element in the state vector is the wage, so \( \theta_w \) parameterizes only the compensation schemes chosen by firms. Section 4.2 presents details about the labor supply choice, while Section 4.3 presents information about state variables and compensation schemes. Section 4.4 discusses the identification of the model. These sections provide methodological details that might not be of interest to all readers, and can be skipped relatively safely.

In order to reduce computational time I estimate the parameters \( \beta \) and \( \theta_w \) separately. This requires that I assume the disturbances in the choice and transition functions are independent. Separate estimation is still consistent, but tends to increase the standard errors relative to joint estimation. I first estimate the compensation schemes, because they are needed as an input into the worker’s decision problem. I present the estimates for the compensation schemes in Section 6 and the results for the workplace choice decision in Section 7.

4.2 Discussion of the Labor Supply Model

This section discusses a worker’s dynamic discrete choice problem of where to work. Previously, in equation 1 I described current period utility as a function of unknown movement costs \( x(j_{i,t-1}, j_{it})' \beta_{xt} \).

---

13I do not compute measures of the productivity benefits of intertemporal relationships because I cannot distinguish effort \( e_i \) from other nonwage attributes of the workplace \( d_i \). By imposing first-order conditions for the optimal choice of \( e_i \) and \( d_i \) on the part of the firm, I could presumably identify the difference between the marginal productivity gain to the firm from worker effort and the marginal cost of providing nonwage aspects of the workplace. However, some nonwage aspects of the workplace include variations in tax policies and living conditions in particular regions that are not under the control of individual firms.
Now I reveal the full list of covariates:

\[ u(j_{it}, s_{it}) + \epsilon_{ijt} = \beta_w w_{ijt} + \beta_{gt} \text{dist}(j_{i,t-1}, j_{it}) + \sum_{l=1}^{3} \beta_{ml,t} m_l(j_{i,t-1}, j_{it}) + \xi_j + \epsilon_{ijt}. \]  

(8)

Here \( w_{ijt} \) is the logarithm of the wage offered by firm \( j_{it} \). The compensation scheme may be very complicated, but each period the worker receives only direct utility from his current wage.\(^{14}\) The following terms allow me to estimate the degree to which labor market frictions impede the movements of workers between firms. The term \( \text{dist}(j_{i,t-1}, j_{it}) \) is the logarithm of the distance in kilometers between a worker’s old firm, \( j_{i,t-1} \), and his new firm \( j_{it} \). Presumably a worker will prefer to not have to move to another region. While I observe the county where each firm is located, I do not observe the residence of each individual worker. Therefore, I assume all workers live in the capital city of the county where their job is located.

The model accounts for the fact that movement induces other penalties beyond geographic moving costs. A worker who changes employers may face an adjustment cost even if he switches to another firm across the street, while a worker internally transferring to a different establishment owned by the same company may face an easier move. Likewise, switching industries may require more adjustment costs. The three \( m_l(j_{i,t-1}, j_{it}) \) terms are indicator variables that equal 1 when, respectively, a worker switches firms, switches between firms/establishments not owned by the same legal company, and switches between firms in different industries.\(^{15}\) I let all of the coefficients on movement costs vary by worker age, as older workers are likely to be tied more closely to existing work and community relationships. Older workers may also own more personal belongings and have more expensive housing, adding extra costs to an employer move that also involves switching residences.

The fixed effect \( \xi_j \) is pivotal to my theoretical interests. It refers to attributes of the work environment that are unobserved in the data. If intertemporal relationships exist, then \( \xi_j \) captures both the costs of putting forth effort and other potentially positive attributes of the workplace, such as the levels of challenge and interest in the work being done. More specifically, \( \xi_j = d_j - e_j \), where \( d_j \) represents the nonwage benefits from a workplace, and \( e_j \) is the required effort level at a particular job. I do not directly observe either of these variables; they vary at the same frequency, and workers treat both of them as fixed when evaluating jobs. It makes sense that, without additional assumptions, I cannot distinguish between unobserved costs and benefits of a job using labor supply information alone. Both the compensation scheme and decision of where to work are chosen knowing \( \xi_j \), thus, I need to account for it in my model. For tractability reasons, I let \( \xi_j \) vary by firm and not with other characteristics such

\(^{14}\)This additive specification for current period utility is typical in the discrete choice literature. A key missing term that is usually included in economic models is total income. A worker with more income from investments might be less interested in wages from work and more interested in nonwage aspects of a job. I exclude income, because, quite simply, I do not observe it in my data. However, the exclusion of total income may not be unreasonable in my context because male, full-time, white collar workers are highly committed to the labor market, up until the age when they choose to retire.

\(^{15}\)The basic unit of analysis is the establishment, and I informally use the term firm to refer to an establishment.
as worker age. However, even with this simplification, I cannot estimate a separate fixed effect for each firm for computational reasons.\footnote{I assume that the $\xi_j$ are known with certainty by all market participants. In other words, the general nonwage amenities and effort intensity at every firm are known by all potential employees.}

The error terms $\epsilon_{ijt}$ are needed to fit the data. In any data set, observably equivalent agents choose different actions.\footnote{In any empirical model, a critical assumption is how the error term is to be interpreted. I will follow the advice of Rust (1994) and view the error term as an unobserved (to the econometrician and to potential employers) state variable that affects preferences. Workers determine the realized value of the vector of unobserved state variables at the start of each period, although firms never learn these values. Under this assumption, the choice probabilities in equation 4 are what is seen by both firms and the econometrician. The error terms are also used in the agent’s calculation of expected future utilities.} For this dynamic programming application, picking the logit distribution for the errors dramatically increases the scope of the models that I can estimate. However, the logit assumption is not without a loss of generality, because it rules out cases where some firm choices are always strictly dominated by others. This is formalized in the following assumption:

**Assumption 1.** Each individual worker picks all employers with positive probability.

This assumption may not be that far from the truth for this data set, as Sweden was a full employment economy during my sample period. Still, I weaken this assumption a little in Section 7.4 by including firm-specific movement costs. I do not estimate a two-sided matching model, because this literature is often advanced by assuming away critical issues that I deal with here, such as the dynamics of labor supply, the endogeneity of wages, and the fact that on the margin firm sizes are not fixed.\footnote{Another problem is the well-known independence from irrelevant alternatives (IIA) property of the logit model, which states that in a static model individual workers’ substitution patterns between different firms are functions of only their choice probabilities, and not the similarity of the different firms to each other. Modeling substitution between employers as compensation policies change is very important, because worker turnover is the same thing as substitution in this dynamic model. As Rust (1994) points out, the IIA property does not exist in dynamic models, because the value functions, $v(s_{it}, j_{it})$, are discounted sums of all the future states a worker may find himself in, including those at other firms. So each $v(s_{it}, j_{it})$ is a function of all the characteristics of the firms operating in the labor market, and the IIA property of the static logit model does not reoccur here. However, to my knowledge, no one has shown that the dynamic logit model is flexible in the sense that it can approximate any arbitrary system of substitution patterns. It may be that this tightly parameterized model puts other, unknown restrictions on substitution and turnover patterns. These unknown restrictions may have econometric benefits, as flexible models may over fit a model to particular data and give results that are hard to interpret and difficult to generalize to other settings. I plan to explore generalizing my empirical model to allow for more flexible distributions of the error terms.}

### 4.3 Compensation Schemes and States

A firm’s major lever to achieve its goals in the labor market is its compensation scheme. A firm offering an attractive compensation scheme can recruit, retain and motivate labor. In this section I formally define what a compensation scheme looks like in the context of my empirical model.\footnote{Swedish white collar workers are generally paid a straight monthly salary, and as a consequence are rarely paid via more complicated schemes such as piece rates. Theoretical explanations for this include the possibility that effort is poorly measured in a legally-binding way and that white collar workers have multiple tasks and rewarding any one of them will cause workers to spend an inefficient amount of time on that task. Unionization may also play a role, as unions may prefer simple pay schemes which give less discretion to management. Whatever the reason, I assume that workers earn only monthly salaries.}

I have discussed how compensation schemes can be subsumed into the theoretical notion of the transition probabilities for a worker’s state variable, $s_{it}$. The three variables that enter the worker’s
vector of state variables, $s_{i,t}$, are his age, current employer and wage. It is important to keep track of a worker’s age because workers tend to receive higher wages and more prestigious assignments as they age: a worker’s utility is nonstationary. A young worker faces a longer time horizon, as he has more time before he retires. Both workers and firms may be more willing to make investments in their relationships if a worker is young, because there will be more time to earn returns from those investments. The current employer is also an important state variable, because the employer identifies the compensation scheme facing a worker. In equation 8 I model utility from accepting a new job as depending on both geographic and non-geographic moving costs. According to my assumptions, knowing where a worker is employed determines the distance and potential moving costs to any other firm.

A worker’s wage is also a state variable. A worker at a particular firm who earns more money than other workers of his same age is probably a more successful worker. That worker has a better ability, has responded better to incentive schemes, has accumulated more human capital, has been lucky, or is doing well for some other reason. My data do not have measures of these potential underlying reasons for worker success, however I somehow need to account for the heterogeneous positions of workers. My measure of the relative position of a worker is the wage of the worker, considered relative to others of his same age at the same firm. Within a given firm it is reasonable that there is a monotonic transformation from the unobserved reasons for success to the observed outcome variable, wage. The actual values of the transformation are not important for the economic questions I address in this paper.

I discretize both a worker’s age and his wage for later estimation purposes. For choice purposes in the labor supply model, I use eight five year age categories (25-29, 30-34, 35-39,...) and ten wage categories for each of the eight age groups. I transform each wage into the median wage in its corresponding age-specific wage decile (i.e., the 5th, 15th, 25th,... quantiles of the age-specific wage distribution).

The wage that enters an age $t$ worker’s state variable, $s_{i,t}$, is the lagged wage $w_{i,j,t-1}$. This means that a worker does not learn the level of his raise for the current period until he has committed to remain at his current firm, and he does not learn the wage he will receive at a new firm until he accepts an offer there. He does know the distribution of wage offers at each firm when making his decision, and this distribution is a function of his past wage. I make this timing assumption about the revelation of the exact wage because I do not observe the wage offer the worker receives at any of the firms a worker declines to choose in period $t$. The lagged wage is a state variable that I can always observe and does not asymmetrically treat observed and counterfactual choices.
In order to account for the fact that a worker only knows what he earned last period, \( w_{i,t-1} \), I redefine the choice specific value functions, equation 3, to be

\[
v(s_{it}, j_{it}) = E[u(j_{it}, s_{it}) \mid s_{it}, j_{it}] + \delta \sum_{s_{i,t+1} \in S} V(s_{i,t+1} \mid s_{it}, j_{it}, \theta_w) = \beta_w E[w_{it} \mid s_{it}, j_{it}] + \beta_g \text{dist}(j_{it-1}, j_{it}) + \sum_{l=1}^{3} \beta_m m_l(j_{it-1}, j_{it}) + \xi_j + \delta \sum_{s_{i,t+1} \in S} V(s_{i,t+1} \mid s_{it}, j_{it}, \theta_w),
\]

where the second equality comes from substituting equation 8 for \( u(\cdot) \). My timing assumption is internally consistent, but may seem restrictive. Again, it is motivated by my need to account for unobserved heterogeneity in the abilities for workers. I view the past wage as a sufficient statistic for the past workplace successes of a worker, or at least the worker’s success relative to others of his same age at his firm.

### 4.4 Overall Identification Strategy

#### 4.4.1 Relation to the Simultaneous Equations Literature

Identification of labor supply curves is difficult. The problems of identifying labor supply curves without individual data can be demonstrated in a simultaneous equations framework. Consider a rudimentary, static version of the labor supply and demand model that I presented in Section 3. The main dependent variables are the quantity of labor at each firm, \( L_j \), and some fixed wage for all workers at each firm, \( w_j \). Workers still care about the unobserved workplace amenities captured by \( \xi_j \), and firms still have unobserved technology parameters \( \omega_j \). For a labor market with \( J \) total firms, this setup results in a simultaneous system of \( 2 \cdot J \) equations that are, for each firm \( j \),

\[
L^S_j = LS_j(w_1, \ldots, w_J, \xi_1, \ldots, \xi_J)
\]

\[
L^D_j = LD_j(w_1, \ldots, w_J, \xi_1, \ldots, \xi_J, \omega_j).
\]

Again, I have dramatically simplified the labor supply (\( LS \)) and labor demand (\( LD \)) curves. The labor supply decision of workers has no dynamics and is based only the observed wages and unobserved could be considered part of the unobserved states \( \epsilon_{it} \). The reason I do not do this is that the wage from the chosen firm would then influence the distribution of unobserved state variables because the lagged wage is a state variable, which means that the time path of future states for the rest of the worker’s life would now depend directly on \( \epsilon_{it} \). Entire lifetime state paths would have to be recalculated for each value of the vector \( \epsilon_{it} \), which is computationally prohibitive. In more technical language, this alternative approach of integrating out unobserved wage offers violates the conditional independence assumption of Rust (1987), which is, by reason of computational necessity, used in all large-dimensional dynamic discrete choice problems. Rust’s assumption, which I adopt, lets the unobserved states \( \epsilon_{ip} \) affect future utility only through the current period’s choice of where to work, \( j_{it} \). As a further point, solution to the integral under the alternative suggestion would not be analytic and would be of an extremely high dimension.
workplace attributes. The labor demand decision of firms also has no dynamics, yet it is still based upon the actions of other firms. Firms offer only one wage, instead of a richly specified compensation scheme.

This is a differentiated-products model of labor demand and supply. Each firm has different attributes $\xi_j$ and $\omega_j$ which cause it to adopt different wage policies. Benkard and Berry (2002) show that even this restrictive model is not identified using aggregate data. With $J$ included endogenous variables $w_j$ in each equation, I need $J$ excluded variables that are included in other equations for instruments. The data do not contain such variables, and so it is clear that I cannot estimate even this simplified model using aggregate data alone.

I solve this identification problem by estimating labor supply curves using data on individual workers. I address the problem of the correlation between the choices of compensation schemes and unobserved effort and work environments by estimating the choice-specific fixed effects $\xi_j$ as included parameters in a worker’s utility function. The fixed effects can be estimated simultaneously with the other parameters because I possess individual data with covariates that also vary at the individual level. Workers treat the compensation schemes and work environments as fixed and choose the best option from among those presented. Thus my use of individual data on worker decisions allows me to estimate labor supply curves without imposing a specific functional form for labor demand, as long as I am willing to assume that workers are not strategic actors, and that compensation schemes are not functions of the idiosyncratic preferences of workers, $\epsilon_{ijt}$. It does not matter whether or not the fixed effects $\xi_j$ are chosen by the firm (as in Section 3) or are instead exogenous characteristics of the type of work the firm performs. My estimation procedure does not rely on a formal revealed preference argument for firm decisions in order to estimate labor supply curves. This is important in a country like Sweden, where it is not clear whether market forces or political factors such as union negotiations are the true driving forces in wage determination.

### 4.4.2 Relation to the Literature on Worker Turnover

The extra information that my data contain that most others do not is the compensation schemes (and successes at recruiting workers) of the other firms operating in the labor market. Varying the number of competitors and their chosen compensation schemes shifts the labor supply curve facing individual firms. This provides identifying information as to the effectiveness of particular compensation schemes in given situations. For the purposes of relating this paper to other previous on turnover where data on the other firms in the labor market are not available, the extra information on outside options can be loosely considered in an instrumental variables framework. Take for example a simple model where worker turnover is a function of only the compensation scheme at the current firm, unobserved variables like nonwage workplace attributes and the intensity of work effort will affect both the decision to leave the firm and the compensation scheme offered by the firm. Given that the compensation scheme is chosen considering the equilibrium compensation schemes of the competitor firms, I can then use instru-
mental variables methods to derive an estimate of the effect of the current firm’s compensation scheme on turnover. This intuition has been operationalized by Bingley and Westergaard-Nielsen (2002), who use the mean of competitor firms’ personnel policies as instruments in a linear regression of a firm’s profits on its own personnel policies.

The drawback of the instrumental variables approach is that the necessary exclusion restriction is not formally valid: the compensation schemes of competitors directly affect turnover and cannot be excluded from the turnover equation. Instead of explicitly using an instrumental variables framework, I implement a discrete choice method in which a worker evaluates the total utility from every employer, and picks the firm that maximizes his utility. After estimating the parameters of a worker’s utility function, I can construct labor supply curves and explore counterfactual changes in the environment facing the worker in order to calculate the sensitivity of the worker’s turnover decisions to changes in the compensation scheme at his current firm.

4.4.3 Worker Utility Identification and Data

Aguirregabiria (2002) shows that the specific choice of the functional form for current period utility in equation 8 is unimportant for identification purposes, because the current utilities $u(j_{it}, s_{it})$ are non-parametrically identified, subject to a normalization. He builds on earlier work by Magnac and Thesmar (2002), who demonstrate that the utility function identification requires that one assumes a particular value for the discount factor and a known distribution for the error terms.\(^{24}\)

I estimate the model using a cross-section of workers who are observed in both 1983 and 1988, which is a period of relative stability in the Swedish manufacturing, white collar labor market.\(^{25}\) Let individual workers be indexed by $i$. For each worker, I know his state variable in 1988, $s_{it}$, which is a vector of the firm $j_{i,t-1}$ he worked at in 1983, his wage $w_{i,j_{i,t-1},t-1}$ at that firm in 1983, and his age in 1988, $t$. At some point in between 1983 and 1988, each worker chooses a firm $j_{it}$ to work at, and later finds out his new wage, $w_{it}$. These last pieces of information form the new state variable for the choice of where to work between 1988 and 1993, $s_{i,t+1}$. Because the state variable includes the lagged wage, the lagged firm choice, and the deterministically increasing worker age, I do not actually need to observe the worker in 1993. All other data that enter the model are indexed to firm identity. These include firm location, industry, and ownership structures for establishments that are all owned by the same national company. Because I use an extremely homogeneous sample of full-time male workers with the same educational background from Sweden, I do not include any other individual covariates that do not vary with firm identity. Of course I have individual data in that I see workers with different ages, wages and geographic

\(^{24}\)It is possible to identify the discount factor through an exclusion restriction that shifts the probability of future states without affecting current period utilities. Also, Taber (2000) shows that an overwhelming number of exclusion restrictions are needed to distinguish terminal period utility functions from distributions over the error terms that allow for across-worker heterogeneity.

\(^{25}\)A particular blue collar labor union broke with the national wage bargaining system in 1983, but that event did not seem to have major impacts on the institutional framework for white collar workers during the sample period.
locations.

Because I only need to observe two years of data, all the identification comes from cross sectional variation in the data. For each firm I observe workers in different stages of their careers. The key identifying assumption is as follows:

**Assumption 2.** A worker of age $t$ believes his world in $\tau$ years will be just like that experienced by workers of age $t + \tau$ today.

I assume that workers do not predict how the macro environment changes. Instead, they look only to the experiences of older workers for information on their own future outcomes. Workers assume the same labor market equilibrium will continue on forever. To be very clear, workers are uncertain about their own individual prospects, because compensation schemes are probability distributions over future wages. Workers are not unsure about what the compensation schemes are or how they will change.

Of course, in the real Swedish economy, the macro environment is constantly changing. The government occasionally undertook currency devaluations in order to improve Sweden’s export competitiveness. These devaluations seriously eroded the real value of wages during the sample period. An offsetting effect is that income taxes have been on a general downward trend. A serious recession hit Sweden just after the sample period, in 1991, perhaps related to Iraq’s invasion of Kuwait. This was the first time that Sweden had a serious unemployment problem, and it is likely that workers did not fathom that such an extreme downturn could happen. The labor market itself is constantly changing, as larger and larger numbers of young workers are earning university degrees, firms are entering and exiting and different sectors are growing and shrinking. It is questionable whether workers can predict these changes, and it is even more unlikely that workers can predict the effects that macro changes have on the labor market. Therefore, given that this paper is about the effects of labor market competition on compensation schemes, I am choosing to ignore macro changes by imposing the strong assumption that workers cannot predict them.

26 A major exception might be the chances a particular firm could fail and lay off its workers. It is clear that small Swedish firms are much more likely to close, and I am sure workers take that into consideration when accepting offers of employment. 27 I have a 21 year panel of Swedish workers. It seems unusual to not use the panel aspect of the data more. Assuming that workers are uncertain about future developments in the overall labor market is computationally prohibitive, because workers have to predict alternative labor market equilibria in counterfactual states. Even if I could estimate this uncertainty model, the identification of the effects of interest will still derive primarily from the cross section, as I do not observe the entire lifetimes of individual workers. Overall, the uncertainty model is probably identified but is not practical. An alternative assumption is that workers know future labor market developments with certainty, and each year has its own set of states. This model is not identified. I do not observe data from 1993, so I cannot estimate the value functions for that year. The value functions for 1993 are inputs into the decision for 1988, so I cannot estimate that year’s value functions either. Backward induction implies that nothing is identified. These arguments indicate that the major source of identification in my data is cross sectional, regardless of how many years of data I use. Clearly, the panel aspect of the data is useful for identifying additional economic variables. The panel data could be used to address questions of worker sorting by estimating types of workers who have different tastes or abilities. I leave this for future work.
5 Data

5.1 Overview

The data come from the Swedish Employers’ Federation, an organization that represents private sector companies in national level wage bargaining with unions. During the sample period, 1970-1990, there were three levels of wage bargaining: national, industry, and establishment level.\(^\text{28}\) In this paper I focus solely on white collar workers, because several empirical facts that motivate my thinking about intertemporal relationships pertain only to white collar workers.\(^\text{29}\) Furthermore, in Sweden white collar unions were pursuing fewer social and political objectives, and thus white collar wages were based more on merit. In this paper, I concentrate on the outcomes of workers with a degree from a vocational high school and workers with degrees in engineering from elite technical institutes.

The data cover nearly 60% of the Swedish private-sector white collar work force.\(^\text{30}\) The data track workers as they move from company to company in the private sector, but cannot track workers if they stop working or switch to an uncovered sector.\(^\text{31}\) The data have information on the salary and job assignment of workers, and on the location and industry of firms. The data do not contain information on the personal characteristics of workers, such as their family status, or on the financial performance of firms.

I view a labor market as composed of firms who hire workers with particular skills. With this data, it is easiest to view workers with varying educational backgrounds as being participants in different labor markets. Most Swedish workers complete their schooling before entering the labor market, so education can be viewed as an unchanging background variable. It is also likely that those workers with business degrees perform different tasks than those with technical degrees. An unfortunate quirk of the SAF data is that only about half of the firms report the educational background of their workers. The firms who do report tend to be the ones most concerned about education, and those firms tend to hire many more technical workers on average than other firms. Therefore, looking at technical workers comes closer to the ideal of observing all workers in Sweden with a particular degree.

Because I want to study the interaction of external labor markets and intertemporal relationships for a particular type of worker, I will consider the data for each educational group separately. I look at two groups of technical workers: those with a three year technical degree from a vocational high school, and those with an engineering degree from an elite university equivalent. A third of technical high school graduates work in the private sector, compared to four-fifths of university engineers. The

\(^{28}\)There was actually lots of wage drift, which is wage increases negotiated at the local level.

\(^{29}\)For example, in Fox (2002) I show that larger firms give higher percentage wage increases over a career where a worker stays at one firm. This is true for American and Swedish white collar workers, but is not true for blue collar workers from either country. It is possible that intertemporal relationships are not as important for blue collar workers.

\(^{30}\)I calculate this based mainly on numbers in Calmfors and Forslund (1990). The main exceptions are the banking industry, which is represented by a different employers’ federation, firms which are cooperatively owned by their workers, and firms which do not belong to any employers’ federation. The data coverage is much better for workers in manufacturing.

\(^{31}\)There is no particular age pattern for workers who leave the data, except for the expected spike at retirement. 
work undertaken by these two groups may also be more tied to their schooling than that performed by non-technical workers.\footnote{I lose track of workers who move to a firm outside of the sector covered by the Employers’ Federation, who leave the labor market, who are promoted to executive status, or who leave the country.}

I utilize age as my measure of a worker’s career stage, because I do not directly observe total experience in the labor market or seniority on a specific job. For workers who do not leave the labor market or become unemployed, age is an excellent proxy for total labor market experience. During the sample period, Sweden was a full employment economy with unemployment levels around 1%. However, in this same time period many workers participated in training programs or were offered early retirement plans. Anecdotal evidence suggests that the more highly educated workers considered here were among the least likely workers to become unemployed. Also, I consider only male workers due to the fact that female workers often temporarily leave the labor market. Next, for part time workers age is a poor proxy for total experience, so I only consider workers whose contractual work time is at least 35 hours a week. Finally, all of the empirical work uses a sample of workers who remain within the data for at least five years and do not start out at a company which goes out of business or move to a company which just opened. In this paper I am trying to address issues facing typical companies and workers, and I do not want to confuse the forces facing these companies with those related to business failure or growing pains. My resulting data set is composed of two 21 year panels of workers with good job prospects and high commitment to the labor market.

In this paper I use the term \textit{firm} as a generic term to refer to any employer. Throughout most of this paper the basic unit of analysis is actually a physical establishment, which I sometimes still refer to as a firm for simplicity. I use the term \textit{company} when I need to explicitly deal with issues surrounding a legal organization that owns more than one establishment. Transfers between establishments owned by the same company happen fairly frequently, so I include a term in the discrete choice model to account for this. The maintained assumption in my model is that all transfers between establishments owned by the same company are voluntary.\footnote{For engineers, 21\% of all observed moves are to an establishment owned by the same legal company.}

\section*{5.2 Firm Estimation Sample}

In my sample, 95\% of the establishments employing engineers do not meet the criteria for estimating compensation schemes using only within-firm variation. The most common problem is that there are very few workers at an establishment and there is not someone who stayed at the establishment for five years (between 1983 and 1988) in a particular age category. For the majority of establishments which do not meet the criteria, I pool them into larger groups based upon region. I refer to the pooled establishments as fringe firms, because most of the establishments in the pooled groups are very small. These estimation groups will be used only to estimate parameters; in the choice model below workers will be able to choose from all firms, or at least a selection of real-life firms, as opposed to pooled groups...
of firms. Turnover between firms in the pooled categories is important, therefore I use pooling only for the necessary steps of estimating parameters of the compensation scheme and, in the labor supply model, for estimating fixed effects.\footnote{When I pool for the purposes of estimating compensation schemes, I still define movers and stayers using real-life establishments, and not the pool definitions.}

Table 3 lists some sample statistics about firms in my estimation sample. For engineers, there are 1039 total firms, which, for estimation purposes, I condense 989 of them into 20 pooled groups based upon their region. 50 of the firms are large enough to estimate the compensation schemes using only within-firm variation. Many of the firms are in Stockholm, and the two largest industries are “manufacturing” and “industry and chemicals.” Unfortunately, dealing with 1039 choices in a nonlinear discrete choice model is computationally intensive. In order to increase the speed of my estimation procedure, I narrow the sample by eliminating firms with less than 4 workers in either 1983 or 1988 for engineers, and less than 7 workers in either of those years for vocational workers. I call these the reduced samples in Table 3. From now on, all results are generated from the reduced sample, although I have verified that the easier-to-calculate implications extend to the full sample.

### 5.3 Worker Estimation Sample

Table 4 lists descriptive statistics for individual workers from both educational groups that I study. While I have data for other years, in this paper the majority of the results are for the five year period 1983-1988. Therefore, I concentrate on that period in the table. Monthly salaries are in Swedish crowns normalized to their value in 1988.\footnote{Because the goal of this paper is to study the relative positions of workers in the economy, I normalize wages so that the aggregate wage distribution (for a particular educational group) has the same mean and standard deviation across all years.} The five year growth in salaries reflects the average growth in the normalized wage levels. The five year movement rate reflects the percentage of workers who move to another firm, with both the new and old firms remaining in business over the period. The variable densely populated county is an indicator variable for locating in one of the three largest metropolitan areas in Sweden. Establishment size is the number of other white collar workers at a worker’s place of employment, while the number of similar employees is the number of employees of the worker’s educational background at that establishment.\footnote{The estimation sample does not include all employees of identical educational backgrounds, as I deleted workers if they are part time, female, or if before 1988 they move to firm not in the sample in 1983.} From Table 4 it appears that engineers are paid more and receive higher percentage raises than graduates of vocational high schools. Engineers also move between firms more often, and are more likely to live in big cities. Finally, engineers tend to work in larger establishments.

### 5.4 Worker Turnover Across Estimation Groups

The main outcome variable I am trying to explain with my structural model is worker turnover. In Figures 1 and 2 I present count histograms of the empirical probability of switching firms for workers
Table 3: Facts About Firms/Establishments in the Estimation Sample\textsuperscript{a}

<table>
<thead>
<tr>
<th>Variable</th>
<th>Vocational</th>
<th>Engineers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full</td>
<td>Reduced</td>
</tr>
<tr>
<td># of Firms (Establishments)</td>
<td>2376</td>
<td>450</td>
</tr>
<tr>
<td>Minimum Size for Inclusion in Reduced Sample</td>
<td>N/A</td>
<td>7</td>
</tr>
<tr>
<td># of Firms with &gt; 0 Obs in 1988\textsuperscript{b}</td>
<td>2375</td>
<td>450</td>
</tr>
<tr>
<td></td>
<td>Remaining for &gt; 0 Obs Firms in 1988</td>
<td></td>
</tr>
<tr>
<td># of Pooled Groups</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td># of non-Pooled Firms</td>
<td>63</td>
<td>49</td>
</tr>
<tr>
<td>Total Estimation Groups\textsuperscript{c}</td>
<td>88</td>
<td>73</td>
</tr>
<tr>
<td>Average Obs. in All Firms</td>
<td>6.5</td>
<td>24.7</td>
</tr>
<tr>
<td>Average Obs. in Pooled Firms</td>
<td>4.1</td>
<td>13.9</td>
</tr>
<tr>
<td>Average Obs. in non-Pooled Firms</td>
<td>94.5</td>
<td>113.4</td>
</tr>
<tr>
<td>Average Obs. in Pooled Groups</td>
<td>174.5</td>
<td>232.3</td>
</tr>
<tr>
<td># of Firms in Stockholm</td>
<td>408</td>
<td>92</td>
</tr>
<tr>
<td># of Firms in Small Counties\textsuperscript{d}</td>
<td>1452</td>
<td>257</td>
</tr>
<tr>
<td># of Firms in Manufacturing</td>
<td>788</td>
<td>182</td>
</tr>
<tr>
<td># of Firms in Industry and Chemicals</td>
<td>555</td>
<td>109</td>
</tr>
</tbody>
</table>

\textsuperscript{a} For the compensation scheme estimation, I pool firms lacking the right number and types of observations in order to estimate all the parameters of the compensation schemes.

\textsuperscript{b} A few firms only have valid observations in the first period (1983) because all valid workers move away. The firms are still hiring workers of the given type, but none of the workers in 1988 are not in the sample in 1983, or are part-time one of the years, or are female.

\textsuperscript{c} This equals the sum of the two rows above it.

\textsuperscript{d} Small counties exclude Stockholm, Gothenburg or Malmo.
### Table 4: Descriptive Statistics about Education Groups for the 5 Year Period 1983-1988\(^a\)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Vocational</th>
<th></th>
<th>Engineers</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>1983 Monthly Salary in 1988 Crowns</td>
<td>14,800</td>
<td>3940</td>
<td>18,600</td>
<td>5110</td>
</tr>
<tr>
<td>1988 Monthly Salary in 1988 Crowns</td>
<td>16,200</td>
<td>4180</td>
<td>20,600</td>
<td>5060</td>
</tr>
<tr>
<td>5-year Normalized Salary Increase %</td>
<td>9.1</td>
<td>13.1</td>
<td>11.3</td>
<td>15.4</td>
</tr>
<tr>
<td>5-year Movement Rate</td>
<td>7.8</td>
<td>26.8</td>
<td>17.5</td>
<td>38.0</td>
</tr>
<tr>
<td>Age in 1983</td>
<td>39.5</td>
<td>8.34</td>
<td>37.1</td>
<td>8.3</td>
</tr>
<tr>
<td>Densely Populated County in 1988 %</td>
<td>48.3</td>
<td>50.0</td>
<td>68.8</td>
<td>46.3</td>
</tr>
<tr>
<td>Establishment Size in 1983</td>
<td>1080</td>
<td>1480</td>
<td>1230</td>
<td>1540</td>
</tr>
<tr>
<td># of Similar Employees in 1983</td>
<td>213</td>
<td>362</td>
<td>178</td>
<td>231</td>
</tr>
<tr>
<td>(N)</td>
<td>9200</td>
<td></td>
<td>6129</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) This table uses the reduced sample that I later utilize for estimation.
in each estimation group. I break out the results by worker age. I use estimation groups as the unit of observation because most firms have very few observations. Figure 1 presents results for vocational workers, while Figure 2 is for engineers. Together, these figures demonstrate that worker turnover varies across Swedish firms. Some firms have more turnover than others. The figures also reproduce the result that, in general, older workers change employers less frequently.

Figure 1: Histogram Counts of Empirical 5 Year Turnover Probabilities by Estimation Groups and Worker Age in 1988, for Vocational Workers

6 Compensation Schemes

In this section I provide descriptive evidence on the compensation schemes chosen by firms in equilibrium. White collar workers earn monthly salaries and can only increase those salaries by receiving raises or switching firms. For a worker considering his future at his current employer, the relevant notion of a compensation scheme is the probability distribution over future wages for the rest of his career. Certain firms offer steeper age-wage profiles than others. For a worker contemplating a move to another firm, the relevant notion of a compensation scheme is how the worker’s history at his current firm is valued.
Figure 2: Histogram Counts of Empirical 5 Year Turnover Probabilities by Estimation Groups and Worker Age in 1988, for Engineers
by the new firm. For example, a firm with a low level of nonwage attractiveness might pay high wages in order to recruit labor. A worker switching to that firm should expect to receive a higher wage than average.

In this paper, compensation schemes are the stochastic transition rule \( f(w_{it} | w_{i,t-1}, j_{it}, j_{it-1}, t_i) \). A worker’s position in the labor market is defined by his wage relative to others of the same age at the same firm, and the relative attractiveness of the firm. In this section, I use data on the wage changes of workers who stay at the same firm or who switch firms to estimate \( f(w_{it} | w_{i,t-1}, j_{it}, j_{it-1}, t_i) \). Later, I use the compensation schemes as inputs into the estimation of the labor supply problem in Section 7.37 I present the results for firm stayers and firm movers in separate sections.

### 6.1 Compensation Schemes for Firm Stayers

Because there are a limited number of observations per firm, I can only estimate a very tightly parameterized version of a compensation scheme, even if it is formally nonparametrically identified through normal methods. The main use of a compensation scheme is to predict wage increases from a worker’s point of view in the dynamic programming problem. I choose functional forms which are easy to estimate and report, but this is more a matter of taste. As long as a functional form is flexible enough to predict wages accurately, the exact choice is not essential.

The most critical aspects of a firm’s compensation scheme are the level and slope (over time) of wages. For the wages of firm stayers, those workers who remain at a firm over the five year period, I use the linear specification

\[
    w_{ijt} - w_{ij,t-1} = \gamma_{1,it} + \eta_{ijt}, \quad \text{if } t > 30,
\]

where the \( \gamma_{1,it} \)'s are firm-specific parameters that capture the five year percentage raise given to age \( t \) workers. All wages are in logarithms. I do not estimate the wages for workers who are new to the labor market, as the goal is to explain how workers move across firms and I do not observe what new workers were doing before they entered the labor market. In practice, I do not have enough data to estimate \( \gamma_{1,it} \) precisely for every age in a firm. I thus pool the coefficients into several age categories. The categories are workers currently aged 30-34, 35-44, 45-54, and 55-64. To clarify, this compensation equation is estimated once for each firm in the data, using only data for workers who stayed at the same firm over the estimation period, 1983-1988. The dependent variable is the wage change over this five year period. This results in four compensation parameters per firm. I also estimate the variance of the

---

37It may seem strange that I am devoting time to estimating a function which is neither a supply nor a demand curve. Wages are the analogs of prices in product markets, and the estimation of an individual-specific price equation as a first step is rare in studies of product markets, even though in many cases prices do vary at the individual level. For example, car dealers negotiate prices with consumers, and better negotiators may receive better prices. While individual level price variation may be a second-order issue in product markets, in labor markets it is absolutely critical. Workers have different paths through the labor market, and a major issue is how firms choose to reward past histories. This paper emphasizes age-wage profiles, which are tools for price discrimination on the basis of age. All I am doing in this section is recovering the pricing rules that firms in the labor market use.
Table 5: Results of Firm-Level Compensation Scheme Estimations for Firm Stayers by Age in 1988

<table>
<thead>
<tr>
<th>Concept</th>
<th>Vocational</th>
<th>Engineers</th>
<th>Vocational</th>
<th>Engineers</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Yr. Increase Age 30-34</td>
<td>0.207</td>
<td>0.0712</td>
<td>0.235</td>
<td>0.0758</td>
</tr>
<tr>
<td>5 Yr. Increase Age 35-44</td>
<td>0.118</td>
<td>0.0432</td>
<td>0.131</td>
<td>0.0416</td>
</tr>
<tr>
<td>5 Yr. Increase Age 45-54</td>
<td>0.0374</td>
<td>0.0350</td>
<td>0.0143</td>
<td>0.0412</td>
</tr>
<tr>
<td>5 Yr. Increase Age 55-64</td>
<td>-0.0236</td>
<td>0.0368</td>
<td>-0.0597</td>
<td>0.0501</td>
</tr>
<tr>
<td>Var. of Error Term</td>
<td>0.00930</td>
<td>0.00338</td>
<td>0.0120</td>
<td>0.00468</td>
</tr>
<tr>
<td>Number of Firms/Groups</td>
<td>73</td>
<td>48</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\) The information for these regressions comes from 1983 and 1988. This table reports the average results of 1 regression for each estimation group.

\(^{b}\) This is the standard deviation of the estimates across estimation groups, not a measure of the statistical precision of individual coefficients. Generally, the coefficients are very precisely estimated.

\(^{c}\) This column is labeled Concept because I list the concept that the coefficient represents, rather than the actual variable included in the regression. For example, \(\gamma_{1,j}^{45}\) is the coefficient on period \(t-1\)'s wage for workers aged 45-54, but in practice it measures the 5-year rate of wage increase for these workers.

Table 5 separately lists estimates of compensation scheme parameters from the regressions I ran on job stayers for each firm or estimation group. The main result in this table is that wage increases are much higher for young workers. Engineers who in 1988 were between 30 and 34 received a 24% raise over the previous five years, while at the other extreme engineers between the ages of 55 and 64 had their normalized wage reduced by 6%. These numbers, if graphed, would replicate the concave age-wage profiles familiar to labor economists. Overall, compared to vocational workers, engineers receive larger raises in the earlier parts of their careers and lower raises in the latter parts.

I describe the level of across-firm heterogeneity in wage increases in the Swedish labor market in Figures 3 and 4 for vocational workers and engineers, respectively. I create histogram-like charts where the vertical axis is a count of how many firms or estimation groups are in each bin. The horizontal axis is the average level of five year wage increases for workers of a particular firm. I present a separate chart for each age group. In general there appears to be important variation in the wage policies of Swedish firms, although I strongly suspect both the within-firm and between-firm variation in wages is less than in a more diverse labor market, such as that found in the United States.

Table 6 describes how the estimated compensation parameters for vocational workers and engineers vary with county population density, which is my measure of potential labor market competition from Section 2. Each coefficient in this table represents a different regression. I use each of the six linear com-

\(^{38}\)To calculate the probability mass function for the discrete transition matrix \(f\left(w_{it} \mid w_{i,t-1}, j_{it}, j_{it-1}, t_{i}\right), I assume wage increases have a normal distribution and then use the normal cumulative distribution function to assign masses to each discrete wage category. I use ten wage categories for each age group.
Figure 3: Count Histograms of Across-Firm Wage Increase Heterogeneity for Vocational Workers
Figure 4: Count Histograms of Across-Firm Wage Increase Heterogeneity for Engineers
pensation parameters in separate regressions of the parameters on an indicator variable for being located in a densely-populated county. In order to examine a richer picture of a firm’s compensation scheme, I include a measure of the starting wages received by workers aged 24-26. I do this even though starting wages do not enter into the structural model, because for data reasons I model only transitions between firms, not worker entry into the labor market. I repeat each regression using observations from other five year periods in order to improve the statistical precision of the results. Looking at the relationships between county size and the estimated parameters, it is clear workers in large counties receive higher wages when they enter the labor market as well as higher five year wage increases between a starting age of 25-29 and an ending age of 30-34. It is does not appear that potential labor market competition is strongly correlated with the level of raises for older workers. Perhaps there is no correlation because older workers are already settled in intertemporal relationships, and there is not much competition for older workers as a result. In results not reported here, I include as regression controls firm size and an indicator for being in a pooled firm group. The main result that potential labor market competition is primarily correlated with the compensation parameters for younger workers still holds.

The estimated compensation parameters imply that the effects of potential labor market competition (as measured by county population density) seem most dominant in the early stages of a worker’s career. Later, when workers may be more reluctant to switch jobs, potential labor market competition does not seem to produce additional inequalities. Of course, workers in more competitive labor markets are still earning the premiums they gained earlier in their careers; it is just that for older workers percentage wage increases no longer differ by county population density.

### 6.2 Compensation Schemes for Firm Movers

In order to fully specify the compensation schemes of firms, I also need to estimate the policies firms use to assign wages to workers who switch firms. Unfortunately, the data is thin relative to the richness of potential compensation schemes. I do not observe many potential switches, such as a worker who moves from a particular small firm in north Sweden to a new firm far away in south Sweden.\(^{39}\) Also, there are few older job switchers at most firms. For this reason, I estimate a single job switcher regression on all of the movers in the sample. The coefficients in the regression vary by age, but not by firm identity. However, the covariates in the regression do vary by firm identity. I estimate, for firm switchers only,

\[
\bar{w}_{ijt} = \gamma_{2t}\bar{w}_{ijt-1,t}^{\text{pred old}} + \gamma_{3t}\bar{w}_{ijt,t}^{\text{pred new}} + \eta_{ijt}
\]

\(^{39}\)The data from workers who switch firms is also highly selected, because only those workers who accept outside offers are reported. Job switchers are usually those who received the highest outside offers, and are not representative of all workers. One possibility might be to utilize the theoretical labor supply model to correct for selection bias by assuming correlation between the choice and compensation errors, and then estimate the two parts of the model simultaneously. This alternative specification is a generalization of my model that seems conceptually straightforward, but, due to the large number of firm-specific compensation parameters, is impractical given current algorithms and computing technologies.
Table 6: Regression of Compensation Scheme Coefficients for Firm Stayers on an Indicator for Being in a Densely Populated County, by Worker Age in 1988

<table>
<thead>
<tr>
<th>Concept^b</th>
<th>Vocational</th>
<th>Engineers</th>
</tr>
</thead>
<tbody>
<tr>
<td>log Starting Wage (Age 24-26)</td>
<td>0.0267</td>
<td>0.0300</td>
</tr>
<tr>
<td>(Age 24-26)</td>
<td>(0.0160)^c (0.00448)</td>
<td>(0.0145) (0.00385)</td>
</tr>
<tr>
<td>Increase Age 30-34</td>
<td>0.0375</td>
<td>0.0286</td>
</tr>
<tr>
<td>Increase Age 30-34</td>
<td>(0.0168) (0.00340)</td>
<td>(0.0224) (0.00458)</td>
</tr>
<tr>
<td>Increase Age 35-44</td>
<td>0.0117</td>
<td>0.00786</td>
</tr>
<tr>
<td>Increase Age 35-44</td>
<td>(0.0105) (0.00247)</td>
<td>(0.0122) (0.00311)</td>
</tr>
<tr>
<td>Increase Age 45-54</td>
<td>0.00930</td>
<td>-0.00119</td>
</tr>
<tr>
<td>Increase Age 45-54</td>
<td>(0.00848) (0.00244)</td>
<td>(0.0121) (0.00305)</td>
</tr>
<tr>
<td>Increase Age 55-64</td>
<td>0.0153</td>
<td>-0.00197</td>
</tr>
<tr>
<td>Increase Age 55-64</td>
<td>(0.00880) (0.00261)</td>
<td>(0.0148) (0.00357)</td>
</tr>
<tr>
<td>Var. of Error Term</td>
<td>-0.000753</td>
<td>-0.000055</td>
</tr>
<tr>
<td>Var. of Error Term</td>
<td>(0.00082) (0.000207)</td>
<td>(0.00138) (0.00268)</td>
</tr>
<tr>
<td>Firm-Years (N)</td>
<td>73</td>
<td>1063</td>
</tr>
</tbody>
</table>

^a Every coefficient in this table represents a separate regression. The dependent variable is the coefficient from the compensation scheme regressions. The reported coefficient modifies an n indicator variable equal to 1 when the firm or pooled group is located in one of the three large cities. I also include a constant term in the regression.

^b The concept refers to the coefficient from the compensation regressions that I am trying to measure.

^c Standard errors are calculated normally.

^d This column reports regressions where the compensation scheme was estimated for all five year intervals between 1971 and 1990. The way in which firms are assigned to the pooling groups differs across years. When calculating standard errors, I do not account for the fact that the each five year period overlaps adjacent periods. In addition to the big county indicator and a constant term, I include year effects in these regressions.
where \( w_{i,j,t-1,t}^{\text{pred old}} \) is the wage that the within-firm compensation schemes (equation 9) from the old firm predicts for the worker if he instead stays, and \( w_{i,j,t}^{\text{pred new}} \) is the predicted wage of a generic worker of age \( t \) who has stayed at the new firm over the last five years.\(^{40}\) The idea is to see whether the wages of firm movers are matched to what the old firm would have paid, or if they are similar to what the new firm pays workers of the same age.

This approach is not entirely consistent with the previous literature on distinguishing the returns to seniority on a job from the returns to total experience (or in my case, age) in the labor market. (Topel, 1991; Altonji and Williams, 1998; Abowd and Kang, 2001) The difference between seniority and experience is that seniority resets to 0 when a worker starts at a new employer. The previous empirical papers purport to be able to use this framework to determine whether firms use productivity-enhancing intertemporal relationships that are specific to particular employment matches. All identification in these models stems from wage changes when workers switch firms. I have shown that different firms use different compensation schemes; in particular, age-wage profiles vary across firms. When a worker switches firms, much of his wage change may be due to the differences in compensation policies between the two firms, not the differences between the returns to seniority and the returns to experience. In my specification, the wages of movers are placed above or below the average level for stayers at that company, and then afterwards the profiles follow the same slope. A test for a seniority effect in my model is to see whether a new worker is placed above or below the profiles of existing workers of the same age. However, as a preview of my results, I do not find much evidence for a seniority effect in wages. Instead of looking at prices, another idea to address the underlying question of whether firms use productivity-enhancing intertemporal relationships is to look at quantities, by which I mean worker turnover. Firms with high turnover are unlikely to be using internal labor market arrangements, such as human capital investments or dynamic incentive schemes.

Table 7 lists the results of the regression of the wages of movers on predicted wages using the old and new firms’ compensation schemes. I let this effect vary by worker age, and the main result is that the wage of an older mover seems to be based much more on the compensation scheme of his old employer. For example, 70% of the wage of a younger engineer who moves is based upon his predicted wage at his old firm, while 99% of the wage of an older engineer who moves is based upon his predicted wage at his old employer. While some part of the wages of a firm mover is due to the unique compensation scheme of the new firm, the majority of the wage is based upon the worker’s past history in the labor market. Clearly, information is revealed about a worker’s ability that cannot be erased by switching firms. This information accumulates as a worker ages; by the end of a worker’s life his wage is very specific to him and not subject to much competition on the market. Therefore, the window of higher wage opportunities from switching employers seems to be open only for younger workers.

\(^{40}\)More explicitly, \( w_{i,j,t-1,t}^{\text{pred old}} = \gamma_{t,t} + w_{i,j,t-1,t} \), where again the wages are in logarithms. The term \( w_{i,j,t}^{\text{pred new}} \) for a worker in a particular age category is the average wage of the 5-year job stayers of the same age group at the new firm. I experimented with several different specifications for predicted wages at the old firm and the new firm, and the economic implications are
Table 7: Results of Compensation Scheme Estimation for Firm Movers, by Worker Age in 1988

| Age Group in 1988 | Vocational | | Vocational | | Engineers | | Engineers |
|------------------|------------|----------|------------|----------|------------|----------|
|                   | % Predicted | % Predicted | % Predicted | % Predicted |
|                   | Old Firm    | New Firm  | Old Firm    | New Firm  |
| Age 30-34         | 0.515       | 0.460     | 0.693       | 0.306     |
|                   | (0.0945)    | (0.0942)  | (0.0628)    | (0.0627)  |
| Age 35-44         | 0.899       | 0.105     | 0.734       | 0.270     |
|                   | (0.0402)    | (0.0401)  | (0.0323)    | (0.0322)  |
| Age 45-54         | 1.01        | -0.00361  | 0.893       | 0.113     |
|                   | (0.0393)    | (0.0393)  | (0.0522)    | (0.0523)  |
| Age 55-64         | 1.05        | -0.0522   | 0.991       | 0.00917   |
|                   | (0.0967)    | (0.0967)  | (0.0813)    | (0.0808)  |
| Var. of Error Term| 0.0167      |           | 0.0180      |           |
| Number of Movers (N) | 714        |           | 1077        |           |

* The information for these regressions comes from 1983 and 1988. This table reports the results of one regression for each educational group. The sample is all firm movers from that educational group. Standard errors are in parentheses. The dependent variable is the wage of the mover. The definition of the measures of the predicted wages are in Footnote 40 to the main text.
7 Results for Labor Supply

One piece of evidence that will help clarify whether labor market competition affects intertemporal employment relationships is the marginal effect of turnover with respect to changes in a compensation scheme. If turnover is very sensitive to changes in a compensation scheme, then the firm can cost-effectively reduce turnover and protect intertemporal relationships against external labor market competition. On the other hand, if very large changes in compensation schemes are needed to reduce turnover, then firms cannot cost-effectively reduce turnover, and potential labor market competition has the ability to constrain intertemporal relationships from operating effectively. In this section I present results that indicate whether turnover is sensitive to wages, and if so which age groups are most sensitive to turnover.

As a demand elasticity is an output from the estimation of a consumer demand model, a turnover or recruitment elasticity is an output from an estimation of a labor supply model. The theoretical model of labor supply in Section 3.1 presents the decision of where to work as a discrete choice. The model is dynamic because workers in the labor market care not only about their current wage, but also the time path of future wages. While I must calculate the value functions \( V \) using backwards recursion dynamic programming in order to properly account for future utility in the estimation, the goal of the econometric procedure is to estimate the parameters \((\beta, \xi)\) in the current period utility function \( u (s_{it}, j_{it}) \). Once these parameters are known, the probabilities of choosing particular jobs can be re-calculated for changes in the economic environment facing a worker.

These utility parameters \((\beta, \xi)\) are the underlying primitives of my model of individual worker behavior, and should be, to a large degree, invariant to the specific institutional details present in Sweden. Sweden is known for having less variance in its wage distribution than countries such as the United States. If there is not enough variation in the data to identify a parameter such as the coefficient on wages, this will show up in an imprecise estimate of the coefficient, rather than as an asymptotically biased estimate. Thus it is important to consider the standard errors of the main outputs of the model, the marginal turnover effects that I will explain and calculate, to see if I can statistically distinguish between economically large and small sensitivities of worker turnover to changes in compensation schemes. It turns out that the data is capable of precisely estimating the effects of interest.

7.1 Parameter Estimates

7.1.1 Static vs. Dynamic Estimates for Simple Models

As I discussed in Section 4.4, in a dynamic discrete choice model the discount factor \( \delta \) is not identified. Therefore, I must choose a particular value for the discount factor before I estimate the model. Table 8 examines whether the results are sensitive to this assumption. I list the parameter estimates and generally consistent across specifications.
standard errors estimates for simple models with discount factors set to, alternately, 0.0 and 0.85, using both the vocational and engineering samples. These parameter estimates are for the parameters in the utility function, equation 8 on page 14. The particular value 0.85 over the five year period is equivalent to a discount factor 0.97 per year. Remember that I have normalized wages to a constant level.

The first column in Table 8 for each of the educational groups presents results for a static model, where the discount factor has been set to 0.0. The notes below each table give the exact definitions of each variable. The parameter estimates have sensible interpretations. Wages are positively valued by workers and workers dislike moving long distances. There are very high costs for moving, although the costs are less if the worker stays within his industry or internally transfers within his multi-establishment company. The second column for each group lists the results from a dynamic model where the discount factor is set to 0.85. For engineers, the coefficient on wages is larger in the dynamic model, which is consistent with a story that workers are forward looking, and ignoring this fact understates the sensitivity of worker behavior to wages. However, this story is not corroborated by the results from the vocational workers, because the correlation between the choice of discount factors and the coefficient on wages goes in the opposite direction for that group. Regardless, I believe workers are in truth forward-looking, and from now on I will only consider results from models where the discount factor is set to 0.85.

7.1.2 The Full Model

The specification that I call the full model includes two extra features. First, I let the original four movement costs parameters vary by the seven age groups that choose where to work. Age-specific movements are one plausible explanation for why younger workers are more likely to move. For example, a new worker must move out of his parents’ house or out of a college dormitory anyway, while older workers have families, houses and connections to a workplace and community which are difficult to break. The second addition is that I estimate a fixed effect $\xi_j$ for each estimation group, which represents unobserved attributes of a job that influence employer choices. The fixed effects address the theoretical concerns from the literatures on compensating wage differentials and efficiency wages, which suggest that wages are correlated with these unobserved, nonwage attributes of a job.

Table 9 lists both the key attributes and results from the model. The key finding in this table is that for engineers the wage coefficient is positive and precisely estimated, while for vocational workers the wage coefficient is positive but not statistically distinguishable from 0. Later, I will further argue that it is possible to statistically claim that the wage coefficient for vocational workers is small. I report

41The standard errors come from applying the BHHH approximation to the partial likelihood function of the discrete choice model. It would be best to estimate the standard errors by taking into account estimation error in the compensation schemes, which form the conditional transition probabilities in the labor supply model. However, I estimated the compensation schemes by repeated applications of regressions in Stata. The coefficients from the regressions for firm stayers enter into the estimation for the coefficients for firm movers, so in some sense this is really a three-stage estimation procedure, which causes further complications.

42The relationship between the discount factor and the wage coefficient appears to be monotonic, in that larger discount factors result in larger wage coefficients.
Table 8: Estimates of Five Parameter Labor Supply Models for Different Discount Factors, between 1983 and 1988\(^a\)

<table>
<thead>
<tr>
<th>Parameter(^b)</th>
<th>Vocational</th>
<th></th>
<th>Engineers</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Static Model</td>
<td>Dynamic Model</td>
<td>Static Model</td>
<td>Dynamic Model</td>
</tr>
<tr>
<td>log Wage(^c)</td>
<td>3.68</td>
<td>1.74</td>
<td>1.06</td>
<td>1.52</td>
</tr>
<tr>
<td></td>
<td>(0.801)</td>
<td>(0.115)</td>
<td>(0.439)</td>
<td>(0.0858)</td>
</tr>
<tr>
<td>log Distance(^d)</td>
<td>-0.466</td>
<td>-0.434</td>
<td>-0.371</td>
<td>-0.362</td>
</tr>
<tr>
<td></td>
<td>(0.0113)</td>
<td>(0.0113)</td>
<td>(0.0100)</td>
<td>(0.0102)</td>
</tr>
<tr>
<td>Base Movement Cost(^e)</td>
<td>-3.19</td>
<td>-3.40</td>
<td>-2.68</td>
<td>-3.04</td>
</tr>
<tr>
<td></td>
<td>(0.0409)</td>
<td>(0.0419)</td>
<td>(0.0428)</td>
<td>(0.0429)</td>
</tr>
<tr>
<td>Not Transferring Cost(^f)</td>
<td>-3.26</td>
<td>-3.14</td>
<td>-3.00</td>
<td>-2.77</td>
</tr>
<tr>
<td></td>
<td>(0.0545)</td>
<td>(0.0544)</td>
<td>(0.0503)</td>
<td>(0.0499)</td>
</tr>
<tr>
<td>Switching Cost(^g)</td>
<td>-2.30</td>
<td>-2.30</td>
<td>-1.73</td>
<td>-1.79</td>
</tr>
<tr>
<td>Industries(^h)</td>
<td>(0.0938)</td>
<td>(0.0884)</td>
<td>(0.0671)</td>
<td>(0.0672)</td>
</tr>
<tr>
<td>Estimation Groups</td>
<td>73</td>
<td>73</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>Discount Factor</td>
<td>0.0</td>
<td>0.85</td>
<td>0.0</td>
<td>0.85</td>
</tr>
<tr>
<td>Total Parameters</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-5410</td>
<td>-5408</td>
<td>-7941</td>
<td>-7957</td>
</tr>
<tr>
<td># of Observations</td>
<td>8697</td>
<td>8697</td>
<td>6129</td>
<td>6129</td>
</tr>
<tr>
<td># of Choices</td>
<td>450</td>
<td>450</td>
<td>402</td>
<td>402</td>
</tr>
</tbody>
</table>

\(^a\) Each column represents a separate estimation.

\(^b\) Standard errors are in parentheses, but use only the BHHH approximation applied to the partial likelihood as if it were the true likelihood. In other words, these standard errors do not account for estimation error in the compensation scheme estimations.

\(^c\) Given my timing assumptions, this variable is the expected value of the log of the wage in 1988.

\(^d\) Distance is the number of kilometers between the capital cities of the current firm’s county and the prospective firm’s county. If the firms are in the same county, I set distance to 1.

\(^e\) The base movement cost is accrued whenever a worker moves establishments.

\(^f\) Many movements between establishments are really transfers within the same multi-establishment company. This is an extra movement cost that is paid when the prospective firm is not an establishment owned by his current company.

\(^g\) The worker is charged this cost if his new establishment is in a different industry from his old one.
Table 9: Key Results from the Estimation of the Full Labor Supply Model\(^a\)

<table>
<thead>
<tr>
<th>Parameter Estimate for log Wage(^c)</th>
<th>Vocational</th>
<th>Engineers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.271</td>
<td>0.783</td>
</tr>
<tr>
<td></td>
<td>(0.203)(^b)</td>
<td>(0.174)</td>
</tr>
<tr>
<td>Estimation Groups</td>
<td>73</td>
<td>48</td>
</tr>
<tr>
<td>Discount Factor</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>Total Parameters</td>
<td>101</td>
<td>76</td>
</tr>
<tr>
<td>Movement Cost Params.</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>Fixed Effect Params.</td>
<td>72</td>
<td>47</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-5170</td>
<td>-7459</td>
</tr>
<tr>
<td># of Observations</td>
<td>8697</td>
<td>6129</td>
</tr>
<tr>
<td># of Choices</td>
<td>450</td>
<td>402</td>
</tr>
</tbody>
</table>

\(^a\) Each column represents a separate estimation.

\(^b\) Standard errors are in parentheses, but use only the BHHH approximation applied to the partial likelihood as if it were the true likelihood. In other words, these standard errors do not account for estimation error in the compensation scheme estimations.

\(^c\) Given my timing assumptions, this variable is the expected value of the log of the wage in 1988.

my findings for the other parameters in future sections. I will also construct marginal turnover effects, which offer economically meaningful interpretations of the wage coefficient.

### 7.1.3 Movement Costs

The full model includes four movement cost parameters for each of the seven age groups. Instead of simply listing the 28 movement cost parameters, I list the results of comparisons that describe the economic meaning of the movement costs. In Table 10, I show estimates of the movements costs expressed in multiples of the average one period salary in 1988.\(^43\) For example, the second hypothetical move shows that an engineer aged between 25-29 in 1983 and moving to an establishment across the street

\(^43\)The wage variable that enters the estimation is the expected log of the 1988 monthly wage. Because I enter the wage in logarithmic form, the coefficients are invariant to the wage being multiplied by a constant. These multiples are calculated by a discrete version of the indifference condition that equates marginal utilities of wages and movement costs,

\[
\frac{\partial u(j_t, s_t)}{\partial W} + \frac{\partial u(j_t, s_t)}{\partial MC} = 0.
\]

Let \(W\) refer to the wage that is not logged, or \(W = \exp(w)\). I calculate the one-time wage bonus, \(\Delta W\), that keeps the worker indifferent after a move. This quantity is the solution to the equation

\[
\frac{\Delta u(j_t, s_t)}{\Delta W} \Delta W + \frac{\Delta u(j_t, s_t)}{\Delta MC} \Delta MC = 0.
\]
Table 10: Multiple of Average 1988 Salary Equivalents of Hypothetical Moves, by Selected Starting (1983) Ages$^a$

<table>
<thead>
<tr>
<th>Hypothetical Moves</th>
<th>Vocational</th>
<th>Engineers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age 25-29</td>
<td>Age 35-39</td>
</tr>
<tr>
<td>Change Company</td>
<td>Change Industry Move Distance</td>
<td>Age 25-29</td>
</tr>
<tr>
<td>N N 1 km</td>
<td>11.0</td>
<td>11.2</td>
</tr>
<tr>
<td>Y N 1 km</td>
<td>21.2</td>
<td>23.5</td>
</tr>
<tr>
<td>N Y 1 km</td>
<td>19.3</td>
<td>19.7</td>
</tr>
<tr>
<td>N N 100 km</td>
<td>19.0</td>
<td>18.2</td>
</tr>
<tr>
<td>Y Y 100 km</td>
<td>37.6</td>
<td>38.9</td>
</tr>
</tbody>
</table>

$^a$ These results are derived from comparing the wage coefficient and the estimated movement costs from the full model. I calculate the movement cost using the linearized lower bound, evaluated at the average 1988 salary for workers of that age in 1983. The costs are expressed as multiples of the same average 1988 salary for workers who started at the given age in 1983.

$^b$ There were no industry switchers in the largest age group, so the maximum likelihood estimate of the corresponding cost is negative infinity. In practice I set the corresponding parameter to be a very low number.

and in the same industry but owned by a different legal company would want a five year bonus of 6.2 times his salary.

An important result in Table 10 is that movement costs are higher for older workers. It is well known that older workers move less often, and the fact that older workers are more established in the current lifestyles and work habits is part of the reason why this is the case. In my model, further reasons why older workers move less are that the wages of older movers do not change very much upon a move (see Section 6.2), and older workers by definition have shorter remaining careers to enjoy any wage premium.

Why do the movement costs seem to be so large? First, the movement costs only accrue once, while moving to a firm with a higher wage gives benefits every period in the future. The one-time bonus calculation that I performed does not amortize the movement cost over a career. Second, the data show

Note that when comparing moving to not moving at all, $MC = \Delta MC$. Given the fact that movement costs enter utility linearly while wages enter in as logs, this is approximately

$$\frac{\beta_w}{W} \Delta W + MC = 0,$$

where $\beta_w$ is the coefficient on wage in the utility function. Solving for $\Delta W$, and expressing the result as a ratio to $W$ gives

$$\frac{\Delta W}{W} = -\frac{MC}{\beta_w},$$

which generates the numbers I report in Table 10. These linearized movement costs are lower bounds; the upper bounds are $\exp\left(\frac{\Delta W}{W}\right)$, however the upper bounds are much too large because they ignore the fact that with log wages entering utility, workers prefer to consumption smooth across time.
that the majority of workers do not move in any period, even if they can receive a small wage increase by moving. The empirical model fits these facts by saying that movement costs are high.\textsuperscript{44} High movement costs are consistent with the empirical reality that there is a distribution of wages in offered by firms in the labor market, and workers do not move around instantaneously to arbitrage away these differences. Movement costs have the result of locking in workers to particular jobs, and might cause them to start paying attention to the particular productivity-enhancing incentive schemes offered by their current firm - an issue that I will discuss later. Younger workers are more likely to move because they can capitalize on a different compensation scheme, they have lower movement costs, and they have a longer remaining career to enjoy the gains from a move.

Note that I am not the first economist to suggest that movement costs are very high in the labor market. In a survey of blue collar workers in New Haven, Connecticut, in the late 1940's, Reynolds (1951) finds that 12\% of workers would switch to a new geographic area for a change in permanent wages of under 25\%, 30\% of workers would move for a permanent change in wages of between 25\% and 100\%, and 45\% would not move to a new area under any circumstances. Regarding moves between employers within the same region, Reynolds finds that that 88\% of the manual workers planned to remain at their current jobs. The current wage level was not a predictor for whether workers planned to remain at their current employer, and, even more amazingly, 88\% of war veterans returned to their previous employer after leaving military service.\textsuperscript{45} Reynolds interprets these findings as indicating, among other things, that workers prefer comfortable routines and relationships at existing employers and do not like adjusting to new and uncertain environments.

According to my estimates, why do any older workers move? The combination of the estimates of the mover compensation schemes in Section 6.2 and the age-specific movement costs imply that there is not much financial gain for older workers who move. Any movement I observe is probably due to idiosyncratic reasons, such as a growing frustration with the current employer, a desire to try new things, a desire to be near loved ones, or some other personal reason. There also does not seem to be much labor market competition for older workers, which is a point I will return to later.

\section*{7.1.4 Fixed Effects}

In the full model, I estimate one fixed effect, \( \xi_j \), for each pooled estimation group or non-pooled firm in the sample. I normalize to 0 the fixed effect for the estimation group consisting of fringe firms operating in Stockholm. The fixed effects are important for my theoretical interests because they control for any idiosyncratic factors that may affect workers' decisions.

\textsuperscript{44}It is important to remember that this is a stochastic discrete choice model, therefore increasing the wage of a firm shifts the fraction of workers who will choose the firm. The aggregate worker decisions are not all or nothing. The finding of relatively high movement costs is not a drawback of estimating a discrete choice model with many alternatives, because the coefficients are scaled to be on the order of the \( J \)th order statistic of the logit distribution, when there are \( J \) different employer choices in the model.

\textsuperscript{45}The government mandated that the employers re-hire the veterans, so this experiment is similar to exposing a worker to new experiences and seeing if he will remain at his current job.
unobserved aspects of jobs that affect wages and turnover decisions.

Figures 5 and 6 show count histograms for the different fixed effects for vocational workers and engineers, respectively. The fixed effects are individually very precisely estimated, so estimation error should not be a concern. On the whole, the the average firm has a more attractive work environment than small firms in Stockholm.

![Figure 5: Count Histogram of Firm Fixed Effects for Vocational Workers](image)

![Figure 6: Count Histogram of Firm Fixed Effects for Engineers](image)

In my theoretical model, I specified fixed effects as the difference $\xi_j = d_j - e_j$ between working conditions $d_j$ and effort elicited using an intertemporal relationship $e_j$. In reality, fixed effects capture all sorts of other effects, such as differences in tax rates across locations, differences in the costs of living, and so on.

It might be interesting to know the relationship between the fixed effect estimates and some very coarse firm characteristics. Unfortunately, I do not have very many of both firm characteristics and fixed effects estimates, so this analysis will be brief. Table 11 reports the results from regressions of the fixed
effects on indicators for being a group of fringe firms and being located in a densely populated county, as well as on the continuous variable log of establishment size. The most statistically significant result is that the log of establishment size has a positive coefficient, which means that larger firms have more attractive work environments. Why might large firms be good places to work at, independent of the compensation schemes offered by those firms? Here are some reasons:

1. Large firms are less likely to go out of business, and provide more stability.
2. There is a more diverse set of job possibilities in a large firm.
3. Large firms are very bureaucratic and require little individual work effort.
4. Large firms are industry leaders and the work done there is more exciting or more prestigious.
5. Large firms offer attractive benefits packages.\textsuperscript{46}

Because smaller firms and firms in the pooled estimation groups tend to have higher turnover, this result is not consistent with my hypothesis that stable intertemporal relationships are used to elicit effort that workers dislike. If this were the case, I would expect that there would be a negative relationship between fixed effects and firm size, because large firms with low turnover are able to elicit high amounts of effort, which enters worker utility negatively. Another way of putting this is that high wages for white collar workers in Sweden do not seem to compensate for adverse working conditions.

7.2 Model Fit

This section analyzes whether the model is able to fit the data properly. Strictly speaking, the goal of the estimation of the labor supply model is to calculate turnover marginal effects with respect to changes in compensation schemes, which I discuss in Section 7.3. If these turnover marginal effects are precisely estimated, then extent to which the model fits the data properly is not of central importance. On the other hand, an analysis of the fit of the model can help provide evidence that the model is capable of predicting economic phenomena of interest. I am more comfortable with a model that properly fits the data. In what follows, I use the estimates from the full model, which includes firm fixed effects and age-specific movement costs.

Table 12 lists aggregate turnover probabilities, from both the data and the model predictions. The model probabilities are predicted means taken over the observations in the data. Clearly, at this level of aggregation the model fits the data well.\textsuperscript{47}

\textsuperscript{46}If nonwage benefits count as income for tax purposes, they are included in my measure of the monthly salary, and do not fall into the fixed effects.

\textsuperscript{47}In results not presented, I calculate how well the model matched the turnover patterns of the individual estimation groups. For engineers, the model mispredicts turnover by more than 0.10 for seven estimation groups. For vocational workers, three firms are off by more than 0.10 probability points.
Table 11: Regression of Fixed Effects on Densely Populated County and Fringe Firm Indicators, and Firm Size\textsuperscript{a}

<table>
<thead>
<tr>
<th>Variable</th>
<th>Vocational</th>
<th>Engineers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fringe Firm</td>
<td>-0.101</td>
<td>-0.0414</td>
</tr>
<tr>
<td></td>
<td>(0.0647)</td>
<td>(0.0724)</td>
</tr>
<tr>
<td>Densely Populated County</td>
<td>0.0271</td>
<td>0.0734</td>
</tr>
<tr>
<td></td>
<td>(0.0465)</td>
<td>(0.0594)</td>
</tr>
<tr>
<td>log of Establishment Size\textsuperscript{b}</td>
<td>0.0609</td>
<td>0.145</td>
</tr>
<tr>
<td></td>
<td>(0.0275)</td>
<td>(0.0311)</td>
</tr>
<tr>
<td>\textit{N}</td>
<td>73</td>
<td>48</td>
</tr>
</tbody>
</table>

\textsuperscript{a} The dependent variables in these two regressions are the fixed effects from the model estimates that includes acceptance probabilities and fixed effects. The fixed effects for the estimation group composed of smaller firms in Stockholm are normalized to zero.

\textsuperscript{b} Establishment size is the number of white collar workers at an establishment. For fringe estimation groups, this is the median number of white collar workers among firms in the group.

Table 12: Actual and Predicted Turnover Probabilities\textsuperscript{a}

<table>
<thead>
<tr>
<th>Variable</th>
<th>Vocational Data</th>
<th>Vocational Predicted</th>
<th>Engineers Data</th>
<th>Engineers Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stay</td>
<td>92.2</td>
<td>92.3</td>
<td>0.825</td>
<td>0.826</td>
</tr>
<tr>
<td>Move</td>
<td>7.8</td>
<td>7.8</td>
<td>0.175</td>
<td>0.174</td>
</tr>
</tbody>
</table>

\textsuperscript{a} The actual turnover probabilities use only the sample used in the estimation. The predicted turnover probabilities are the model probabilities summed up over all the data points. The model estimates used are the ones that include acceptance probabilities and fixed effects.
Table 13: Actual and Predicted Turnover Probabilities by Worker Age Category

<table>
<thead>
<tr>
<th>Starting Age</th>
<th>Vocational Data</th>
<th>Predicted</th>
<th>Engineers Data</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-29</td>
<td>0.119</td>
<td>0.119</td>
<td>0.287</td>
<td>0.287</td>
</tr>
<tr>
<td>30-34</td>
<td>0.105</td>
<td>0.104</td>
<td>0.197</td>
<td>0.197</td>
</tr>
<tr>
<td>35-39</td>
<td>0.091</td>
<td>0.090</td>
<td>0.167</td>
<td>0.165</td>
</tr>
<tr>
<td>40-44</td>
<td>0.064</td>
<td>0.064</td>
<td>0.135</td>
<td>0.134</td>
</tr>
<tr>
<td>45-49</td>
<td>0.051</td>
<td>0.051</td>
<td>0.073</td>
<td>0.073</td>
</tr>
<tr>
<td>50-54</td>
<td>0.033</td>
<td>0.032</td>
<td>0.066</td>
<td>0.065</td>
</tr>
<tr>
<td>55-59</td>
<td>0.025</td>
<td>0.024</td>
<td>0.059</td>
<td>0.059</td>
</tr>
</tbody>
</table>

The actual turnover probabilities use only the sample used in the estimation. The predicted turnover probabilities are the model probabilities summed over the data points.

A widely known empirical fact about turnover is that it declines with worker age. (Farber, 1999) In my model, I include several features to explain this fact. The compensation schemes offered by firms vary by worker age. Movement costs also vary by age. Finally, the dynamic programming model with a positive discount factor assumes that younger workers value the wages they will earn over their remaining careers. Table 13 presents measures of turnover probability fits for the different age categories used in the estimation. The model fits the turnover probabilities extremely well.

7.3 Turnover Marginal Effects

The main purpose of the labor supply model is to determine the sensitivity of turnover behavior is to compensation scheme changes offered by each worker’s current firm. If turnover is very sensitive to compensation scheme changes, it is cost-effective to reduce turnover, and any observed turnover must not affect worker productivity very much. The positive sign of the coefficient on wages in the utility function usually generates the result that higher wages reduce turnover, but in a nonlinear discrete choice model the magnitude of the effect is dependent on all of the parameters.

7.3.1 Average Effects for All Firms

The simplest way to measure the effect of a compensation scheme change is to consider what would happen if a firm increases future raises for all of its workers. More specifically, I consider the effects on turnover of increasing all of the parameters $\gamma_{1,jt}$ from the compensation scheme for firm stayers, equation 9, by the same amount. Each of these parameters is a percentage wage increase, so adding a constant to all of them means that workers of all ages will receive an additional percentage wage increase.
increase. While the actual final raise depends upon the age of the worker, the additional amount is the same percentage for all workers. Let the new raises be \( \gamma_{1,jt} + \Delta \), where \( \Delta \) does not depend on age, \( t \). Then the marginal turnover effect for workers of age \( t \) at firm \( j \) is

\[
1 \frac{1}{I_{jt}} \sum_{i=1}^{I_j} \left. \frac{\partial (1 - P(j_{it} \mid j_t, t, w_{it}, \beta))}{\partial \Delta} \right|_{\Delta=0},
\]

where \( 1 - P(j_{it} \mid j_t, t, w_{it}, \beta) \) is the probability of leaving the current firm and \( I_{jt} \) is the number of workers of age category \( t \) in firm \( j \).

For reference purposes, it is important to know what a cost-effective marginal turnover effect is. First, I argue that increasing the rate of five year wage increases by 0.01 is not inexpensive. Table 4 shows that the average average five year wage increase rate for engineers is 0.113. Increasing this amount by 0.01 amounts to about a 10% hike in the amount of the raise. Furthermore, young workers continue to receive larger raises over the remainder of their careers, so the compounding of this compensation scheme change means that a new 25 year old worker will receive a 13% higher base salary at the time he retires. Table 4 also shows that the average turnover rate for engineers is 0.173. In order for my hypothetical compensation scheme change to reduce the average turnover rate by a fourth (by 0.04), the marginal turnover effect must be -4.0. As a rough estimate, I say any marginal turnover effect over -2.0 is cost-effective.\(^{48}\)

Table 14 reports the marginal turnover effects for both vocational workers and engineers. The effects are listed by worker age. Each reported statistic is the across-firm mean of the within-firm average marginal effects. I also include the standard errors of each estimate, which accounts for estimation error. I will discuss the heterogeneity across the marginal turnover effects below. The two main results from Table 14 are that marginal turnover effects are low in absolute value and that they decline rapidly with worker age. For vocational workers, the marginal effects of reducing turnover by adjusting compensation schemes are low for all age groups. The effects are not statistically different from 0, although the lower bounds of the 95% confidence intervals are always greater (lower in absolute value) than -0.5, so for vocational workers I can statistically say that the marginal turnover effects are economically small. The overall magnitude of the effects is somewhat larger for engineers. For the youngest engineers, the marginal turnover effect of -0.856 means that a 0.01 increase in the rate of wage increases for all age groups reduces turnover of the young workers by -0.00856. Table 13 shows that the average turnover probability for aged 25-29 workers is 0.287, so I predict my proposed compensation scheme change will, on average, eliminate 3% of the turnover of young workers. Given that I suggested above that this compensation change is not inexpensive in the long-run, this is evidence consistent with the view that

\(^{48}\) I can report elasticities instead of marginal effects. To compute an elasticity, I take the weighted average of the marginal effects and multiply it by the ratio of the average log-wage increase between 1983 and 1988 and the turnover rate between 1983 and 1988. I find the interpretation of the elasticities confusing, because I have to consider percentages of percentages such as turnover rates and raises. For this reasons, I report only the marginal effects, although readers should know that I can easily transform the results into elasticities that convey the same information.
Table 14: Across-Firm Average Marginal Turnover Effects for Increasing Rate of Wage Increases for all Age Groups

<table>
<thead>
<tr>
<th>Starting Age</th>
<th>Vocational Mean</th>
<th>Vocational Std. Error</th>
<th>Engineers Mean</th>
<th>Engineers Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-29</td>
<td>-0.195</td>
<td>0.146</td>
<td>-0.856</td>
<td>0.187</td>
</tr>
<tr>
<td>30-34</td>
<td>-0.149</td>
<td>0.114</td>
<td>-0.701</td>
<td>0.161</td>
</tr>
<tr>
<td>35-39</td>
<td>-0.081</td>
<td>0.062</td>
<td>-0.403</td>
<td>0.0905</td>
</tr>
<tr>
<td>40-44</td>
<td>-0.047</td>
<td>0.036</td>
<td>-0.250</td>
<td>0.0584</td>
</tr>
<tr>
<td>45-49</td>
<td>-0.023</td>
<td>0.0174</td>
<td>-0.102</td>
<td>0.0237</td>
</tr>
<tr>
<td>50-54</td>
<td>-0.0037</td>
<td>0.0023</td>
<td>-0.023</td>
<td>0.0050</td>
</tr>
<tr>
<td>N</td>
<td>450</td>
<td>402</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a These effects are from equation 11. The standard errors of the effects are calculated using the delta method, where I account for estimation error in only the coefficient on wages in the utility function. This parameter has by far the largest impact on the marginal turnover elasticities. Applying the delta method to account for variation in all of the parameter estimates is computationally prohibitive.

altering compensation schemes is not a cost-effective way to reduce turnover.

Table 14 shows that workers in the earlier part of their career are more sensitive to changes in future compensation schemes than are older workers. For the oldest engineers, the proposed increase in raises of 0.01 for all workers their probability of turning over by 0.00023, or only about 0.3% of their empirical turnover probability. This is a tiny change. If firms engaging in labor market competition are trying to attract workers with compensation schemes, it appears that only workers in the earlier parts of their careers will respond.

Figures 7 and 8 show count histograms of the marginal turnover effects for individual firms. Figure 7 is for vocational workers, while Figure 8 presents results for engineers. It is clear there is a lot of across-firm heterogeneity in the marginal turnover effects.

7.3.2 Turnover Marginal Effects Within and Across Markets

Table 15 lists the results of a regression of the marginal effects on the log of a county’s population density and the log of establishment size. Potential labor market competition, as measured by population density, is negatively correlated with the marginal effects, while establishment size is positively correlated with the effects. It is clear the marginal turnover effects are greater in absolute value in more densely populated counties. However, in economic terms the differences in the marginal turnover effects are not high. For engineers, a firm in a county with twice the population density as another is statistically predicted to have marginal turnover elasticities that is greater by -0.035, a small amount. Compensation
Figure 7: Count Histograms of Marginal Turnover Effects for Individual Firms for Vocational Workers, by Worker Age in 1988

Figure 8: Count Histograms of Marginal Turnover Effects for Individual Firms, for Engineers by Worker Age in 1988
Table 15: Regression of Firm Marginal Turnover Effects on Population Density and Establishment Size$^a$

<table>
<thead>
<tr>
<th>Variable</th>
<th>Vocational</th>
<th>Engineers</th>
</tr>
</thead>
<tbody>
<tr>
<td>log of Population Density</td>
<td>-0.0059</td>
<td>-0.0561</td>
</tr>
<tr>
<td>log of Establishment Size</td>
<td>0.00345</td>
<td>0.0228</td>
</tr>
<tr>
<td>Firm-Age Categories (N)</td>
<td>1935</td>
<td>1469</td>
</tr>
</tbody>
</table>

$^a$ The dependent variable is the marginal effect from equation 11. I also include indicator variables for each age category, as well as a constant term. Standard errors are in parentheses. The population density information is from 1990. Establishment size counts all white collar workers in 1983, although the result reported in the table is robust to counting only workers of the particular educational type, or only counting the fraction of workers in the firm out of all workers in the county (a measure of labor market share).

schemes are not effective ways to reduce turnover in either densely or sparsely populated counties.\textsuperscript{49} Likewise, the statistical correlation of firm size and the estimated marginal effects is not large. For engineers, doubling the size of a firm increases the statistical prediction of the turnover elasticities by 0.013, another small change.

7.4 Robustness: Firm-Specific Movement Costs

One concern with the analysis so far is that I have assumed that workers can pick a job at any firm they choose. In reality, a firm must agree to hire a worker before a new employment relationship can begin. My data does not contain any information on jobs applications that are not accepted, so I cannot easily measure the selectivity of each firm and thus the accuracy of my model where workers choose between all firms. However, I do observe the number of workers who choose to remain at an employer and the number of workers outside of the firm who choose to move there. I can use this information to tell whether a firm is willing to hire a lot of workers. My assumption is that a firm that retains most of its workforce but employs few new workers has few openings. The observed turnover of workers previously at the firm identifies the attractiveness of the firm, and the observed movement to the firm of workers previously outside of the firm identifies the firm’s hiring policy.\textsuperscript{50}

\textsuperscript{49}The equivalent turnover elasticities are also higher in larger counties, so the relationship between county size and the marginal turnover effects is not only due to the fact that turnover is larger there.

\textsuperscript{50}With data on only the workers at one firm and where they choose to work later, it is impossible to distinguish between the attractiveness of the current firm and base movement costs. Seeing the decisions of workers outside of a firm is necessary to separately identify the firm’s attractiveness.
In order to check the robustness of my results, I implement this logic in two different ways. The first method incorporates firm-specific acceptance probabilities into the labor supply model. In this framework, the worker chooses to apply to at most one new firm each period, and has a positive probability of being rejected and forced to remain at his current firm. The problem with this approach is that it introduces extra nonlinearities in the relationships between the various parameters. I feel the nonlinearities obscure the true source of the identification of the acceptance probabilities. To address this extra concern with identification, I estimate another model where each firm has a unique movement cost that is charged to all workers who move to the firm. The results from both models are encouraging in that the main economic conclusions of the paper still hold.

I have estimated many different variations on the basic model presented in this section, and always one of the key results is that the effects of compensation scheme changes on worker turnover are small. To me, the consistency of the results means that minor methodological changes will be less likely to alter the basic economic conclusion that turnover effects are small. These changes, such as letting the marginal effects vary by unobserved worker characteristics, changing functional forms, or changing the composition of the estimation sample, may modify the point estimates of turnover marginal effects somewhat, but given my consistent economic conclusions, I believe future methodological changes are unlikely to produce the result that reducing turnover is cost-effective for most Swedish firms.

8 Discussion of the Results

In this section I first summarize the empirical results in the paper. Then I use the results to create a consistent economic story about firm behavior in the labor market.

8.1 Overview of the Empirical Results

In Section 2, I present empirical evidence that turnover is higher in more densely populated Swedish counties. I interpret this fact as evidence consistent with the idea that potential labor market competition increases turnover. Further, I use the empirical fact as motivation to further study the relationship between turnover and labor market competition. As I will discuss, the empirical evidence from the rest of the paper explains my turnover fact by suggesting that firms are unable to adjust their compensation schemes to effectively reduce turnover in labor markets that are more competitive.

There are six main empirical results from the estimation of the compensation schemes in Section 6 and the labor supply model in Section 7:

1. Potential labor market competition is correlated with wages and rates of wage increase for younger workers but not for older workers.

2. Upon a move, older workers are paid wages that resemble those they would have been predicted
to earn at their old firm, while the wages of younger workers after a move resemble more the compensation scheme of their new firm.

3. Movement costs are large and increase with age.

4. Large firms have attractive workplace environments.

5. The sensitivity of worker turnover to compensation schemes is, in general, low.

6. The turnover behavior of younger workers is somewhat sensitive to a firm’s choice of a compensation scheme, but the turnover behavior of older workers is not at all sensitive to changes in compensation schemes.

The first two results are descriptive results from my analysis of observed compensation schemes, and reflect the equilibrium behavior of firms in the labor market. The last four results are from the structural model of worker behavior in the labor market, and reflect the underlying preferences of workers.

The empirical results describe an interesting picture of how the labor market treats older and younger workers differently. Younger workers are new to the labor market, have fewer social relationships, have less information revealed about them, have fewer firm-specific skills, and have not necessarily committed to a productivity-enhancing intertemporal relationship at any one firm. These workers could potentially be employed by any number of firms and are willing to switch to new employers at a higher rate.

Older workers find it more difficult to move. They have established relationships in the community and their workplace, may have developed specific skills and attachments to interpersonal relationships, and have a shorter future career to enjoy any wage gain from moving. Information has been revealed about these workers that follows them as they change jobs, because their wage is based very closely on the wage I predict they would receive at their old firm. What little turnover that happens for older workers does not seem to be affected by compensation policies, and therefore older worker turnover is mostly caused by variation in personal tastes and idiosyncratic dissatisfaction with general employment conditions at the current employer. A worker may grow tired of a particular employer and wish to move, even if he does not receive much in the way of higher pay. The functional forms for worker turnover, the results about the sensitivity of worker turnover to compensation schemes, and the empirical findings in Section 2 all indicate that firms in less densely populated areas with fewer labor market competitors have less of this idiosyncratic turnover, as do firms with more attractive workplaces.

8.2 Informed Speculation About Firm Behavior in Labor Markets

In my data, I cannot observe the internal operations of firms, but with my results I can make some informed guesses. A clear implication is that worker turnover is not very sensitive to the choice of compensation schemes. This casts doubt on the empirical relevance of theories that emphasize how
firms use compensation schemes to cost-effectively reduce turnover. (Pencavel, 1972; Parsons, 1972; Stiglitz, 1974; Salop and Salop, 1976)

8.2.1 Treatment of Older Workers

It is not correct to suggest that the labor market is perfectly competitive, with workers moving freely in response to offers of small wage increases. My findings suggest that older workers especially face high movement costs and are reluctant to consider outside offers. Older workers who do leave a firm probably do it for non-monetary reasons. Firms apparently respond to this situation by not competing heavily for the services of older workers. For example, in Table 6 I show that the raises given to older workers are not correlated with one of my measures of potential labor market competition, the population density of an area.

From the point of view of the firm, I believe turnover of older workers may almost be thought of as exogenous in the empirically relevant range of compensation schemes. Some firms, with attractive workplaces or few labor market competitors, may have exogenously low turnover and find it more rewarding to invest in productivity-enhancing intertemporal relationships earlier in the their workers’ careers. These investments can be explicit training programs for general or specific skills, or dynamic relationships where workers put forth more effort in return for higher future raises. Of course there could be downsides to the lack of turnover. Low turnover firms might be overly bureaucratic and inefficient. Certain firms have more or less exogenously higher turnover of older workers. These firms find it less beneficial to invest in long-term, intertemporal relationships with their workers.

The previous theoretical literature has emphasized that the amount of worker turnover is a choice of the firm based upon its underlying production technology. I have instead shown that the turnover of older workers is not very sensitive to the choice of compensation schemes. Even if firms do not compete strongly for the services of older workers, it is still the case that some firms have better labor market positions than others, because they have more attractive workplaces or face fewer competitors. One of the questions I posed at the outset is whether or not the availability of more outside options can reduce the effectiveness of productivity enhancing intertemporal relationships. The inability of a firm to reduce its turnover through its compensation scheme indicates that outside options can affect the ability of the firm to develop complex internal labor markets.

8.2.2 Treatment of Younger Workers

I have found several pieces of empirical evidence that are consistent with a story where firms compete more intensely for the services of younger workers. I found that a younger worker is able to take advantages of a new firm’s compensation scheme after a move, I found that younger workers’ wages and wage

51The estimates of the fixed effects, which are greater at larger firms with low turnover, does not seem to indicate a large role for distasteful effort in the decision of where to work.
increases are correlated with my measure of potential labor market competition, and I found that the turnover behavior of younger workers is much more sensitive to the choice of compensation schemes. Aside from the supply-side fact that younger workers are more willing to move in response to higher wage offers, why do firms compete more for the services of younger workers? Again, stories about intertemporal relationships between workers and firms are relevant. If the benefit of a productivity-enhancing relationship is increasing in the amount of time the relationship is expected to last, then it makes sense that firms would want to hire workers as early in their careers as possible.\textsuperscript{52}

8.3 An Explanation for Concave Age-Earnings Profiles?

In Table 5, I show that age-earnings profiles are concave, which means that the wages of workers increase the fastest when they are young. It is interesting that wages increase the fastest when workers supply labor more elastically, and when firms’ equilibrium labor market competition for workers seems to be the strongest. Older workers supply labor less elastically because movement costs are large, and firms seem to be less interested in competing for workers as older workers receive low levels of wage increases. How closely related are facts about labor demand and supply on one hand and the slope of the age-earnings profiles on the other? A model where firms cannot cut nominal wages can generate the result that wages are initially set low but rapidly increase over ages where workers are more elastically supplying labor and firms are more inelastically demanding labor, as firms want to hire younger workers and force them into long-term relationships. Later, firms cease giving large nominal wage increases because they have no need to compete for the services of their older workers, as these workers are locked in because of movement costs.

9 Conclusions

In this paper I have developed an understanding of how labor market competition affects a firm’s choice of compensation scheme. In particular, I find that reducing turnover through compensation scheme adjustments is not cost effective. In general the sensitivity of worker labor supply behavior to wage differentials is low, because workers face high moving costs and are in effect locked into jobs that provide at least a reservation level of satisfaction. Most workers are simply not looking to leave, and firms are also not competing heavily for their services. Because firms cannot cost-effectively reduce turnover, increasing measures of potential labor market competition may have the effect of increasing worker turnover and causing firms to deemphasize the use of internal productivity-enhancing intertemporal relationships. This happens if idiosyncratically dissatisfied workers find new alternative employers that they find more attractive. My empirical findings are inconsistent with a perfectly competitive labor market.

\textsuperscript{52}Hu (1999) presents a similar argument as an explanation for her finding that large American firms tend to hire younger workers. In Fox (2002), I replicate her finding using the more detailed Swedish data.
where workers move for small adjustments in wages, a finding also found in the worker interviews of Reynolds (1951).

The one exception to this finding relates to younger workers; in this case adjusting compensation schemes has a nontrivial effect on worker turnover. Young workers’ responsiveness to wage offers makes it profitable for a firm to engage in labor market competition in order to recruit young workers and place them into long-term careers at the firm. The turnover of older workers, on the other hand, is low because these workers face high movement costs with short future careers to cover these costs. Firms are reluctant to pay older workers large premia to induce turnover, in part because older movers do not fit well into existing internal labor markets. Even though they lack mobility, I presume that older workers are not exploited to a greater degree than they are in reality because younger workers are forward looking and consider their future treatment when deciding where to work.

What do these results have to say about whether or not intertemporal relationships are in fact used to increase worker productivity? Previously the theoretical literature implicitly tied the idea that firms can cost-effectively reduce turnover through adjustments to compensation schemes to the idea that firms use long-term relationships to improve productivity. I have shown than this tying is not logically necessary because older workers face large movement costs are unlikely to leave for small wage differences anyway. Otherwise, all I can say about the existence of productivity-enhancing intertemporal relationships is that the behavior of firms, where in equilibrium they compete more for younger workers, is not inconsistent with the existence of these productivity enhancing relationships. I can add that if intertemporal relationships that enhance worker productivity are important, then the extent to which these schemes can be used is, to some degree, determined by the labor market. Firms with fewer local labor market competitors can make more use of these schemes, because they have lower turnover. Firms in more competitive labor markets will have higher turnover and will have to adjust their internal arrangements to deemphasize intertemporal relationships. If these adjustments, which may include less on-the-job training, fewer internal promotions tournaments, and less extensive human resources departments, reduce worker output, firms’ productivities will also be lower.

The implication of this paper is that policies that increase potential labor market competition or decrease worker movement costs may reduce within-firm worker productivity through increased turnover. This finding must be regarded in light of the benefits of a flexible labor market, which allows the reallocation of labor across firms when shocks to output markets shift the optimal distribution of labor. It may be that the allocative efficiency of competitive labor markets exceeds any productivity gain from increased use of intertemporal relationships. In this case, the deregulation of labor markets can be justified by arguments involving productivity. On the other hand, if intertemporal relationships are widely relied upon by firms, then it could be that a more regulated labor market - such as that which exists in many European countries - increases the overall output level of the economy. This paper moves the debate further by providing evidence that the behavior of firms is consistent with the notion that intertemporal relationships are important, as firms compete more strongly for young workers who are
then locked into life-long careers.

References


