Turnover and Wage Growth for Low Skilled Young Men [1]

Tricia Gladden
Department of Economics
University of Washington-Saint Louis

and Christopher Taber
Department of Economics
University of Wisconsin-Madison

April 7, 2009

[1]A previous draft of this paper was circulated with the title "Turnover and Wage Growth in the Transition from School to Work." We thank Fabien Postel-Vinay, Bob Topel, and seminar participants at a number of institutions for helpful comments. Taber thanks the William T. Grant Foundation for financial support.

1 Introduction

A distinguishing aspect of the early labor force experience of young workers is the amount of turnover that they experience. Their work history is very erratic with many job changes and weak attachment to the labor force. Many argue that this turnover is inefficient and leads to wage stagnation. For example Stern et. al. (1994) states that

> "After leaving school, with or without a diploma, most young people spend a number of years 'floundering' from one job to another, often with occasional spells of unemployment. ... New school-to-work initiatives are intended, in part, to prevent this evident waste of human resources."

On the other hand, others have argued that if the labor force turnover represents job matching then it can lead to wage growth. Heckman (1993) argues that

> "Job shopping promotes wage growth. Turnover is another form of investment not demonstrably less efficient than youth apprenticeships."

The goal of this paper is to quantify the nature of the turnover among young workers. To what extent does this turnover represent "floundering" through jobs and to what extent does it resemble efficient human capital investment?

The basic approach in this paper is to compare the actual path of wage growth to alternative simulated paths for male low skilled workers.[1] We first estimate a behavioral model of turnover and wage growth. We then use the estimates of the model to construct two types of counterfactual wage paths. The first type is designed to quantify the importance of turnover as a component of wage growth. To do this, we simulate the effects on wages and earnings over the first ten years if we could somehow eliminate turnover (or certain types of turnover) for young workers holding all else equal. We experiment with shutting down

[1] We focus exclusively on workers with a high school diploma or less.

various types of transitions including quits, layoffs, job-to-job, and job-to-nonemployment spells. The most important goal of this part of the research is to illustrate the trade-off between the additional wage growth that comes from job matching and the lost wages, experience, and tenure that accompany turnover.

Our results show that the overall effect of turnover on wage growth is clearly positive. Eliminating all turnover substantially reduces wage growth. Perhaps most importantly, eliminating voluntary turnover leads to substantially lower wage growth: wages are approximately 7.3% lower after ten years. Workers seem to be doing more than simply 'floundering' from one job to the next. However, they do not appear to be purely maximizing their earnings either. In fact, eliminating all turnover would increase the present value of earnings during the first ten years in the labor force experience due to the additional employment. One of the most interesting cases is eliminating only voluntary switches to nonemployment. This yields higher wage growth and higher earnings strongly suggesting that workers do not simply maximize the present value of their earnings.

The second type of counterfactual we consider is an "optimal" profile. Given the observed distribution of wage gains in the data, we solve a dynamic programming problem to uncover the pattern of turnover that maximizes the present value of income. We then compare the optimal path of wages to the actual observed path. This allows us to quantify the extent to which turnover is excessive and the extent to which it is important for wage growth.

The "optimal" profile is strikingly similar to the counterfactual in which we eliminate only voluntary switches to unemployment. Once again, restricting turnover increases both wage growth and the present value of earnings - wage growth is about 10% higher when workers are restricted to follow this "optimal" profile. The results are again consistent with the basic finding that workers are not simply floundering from job to job, but are not purely maximizing income either.

Our approach differs from a classical structural approach in two ways. First, we are not sure how to formalize the idea of "floundering." Rather than explicitly specifying functional forms for preferences, we just estimate the choice probabilities conditional on the state variables. Second, it has been known since Flinn and Heckman (1986) that the type of search/matching model we present is not nonparametrically identified as one can typically not identify the full offer distribution. Rather than estimate the wage offer distribution we only estimate the accepted wage distribution (conditional on the state variables). Both of these limitations restrict the types of counterfactual that we can simulate. Nevertheless, we can still perform many interesting simulations.[2] We focus on simulating counterfactual wage paths that can be semiparametrically identified from the version of the model that we estimate. In doing so we use a fixed effect approach following the style used in the literature on estimating the returns to tenure (e.g. Abraham and Farber, 1987, Altonji and Shakato 1987, or Topel 1991).

This work is intended to be positive as opposed to normative. A worker may quit a high paying job to enter nonemployment for a legitimate reason such as to care for a sick child. He also may quit for a less socially desirable reason such as to participate in the use of illegal narcotics. We can not distinguish between these activities with this data. Even if this distinction were possible, we could never reject the notion that agents are acting rationally in that their choices make them better off than if they were simply maximizing income. However, we can and do quantify the difference between their observed behavior and pure income maximization.

The paper proceeds as follows. We begin by describing the previous literature. We then develop an economic framework to think about the problem. Next we describe the data. We then discuss the econometric specification that we use and the parameter estimates that we obtain. Finally, we use the estimates of our behavioral model to simulate the counterfactuals.

[2]For example Ichimura and Taber (2001) present a set of conditions in which one can estimate policy counterfactuals directly even if the full model is not identified.

2 Previous Research

A substantial literature has arisen that examines the effects of job mobility (or job stability) on earnings using either regression analysis or structural modeling. Examples include Mincer and Jovanovich (1981), Flinn (1986), Antel (1991), Loprest (1992), Topel and Ward (1992), Wolpin (1992), Farber (1994), and Klerman and Karoly (1994). Devine and Kiefer (1991) and Wolpin (1995) provide surveys of different aspects of these literatures. All of these papers use very different techniques than this one and none of them study the question of interest here.

A few papers decompose wage growth into different components. Topel and Ward (1992) show that 33% of early labor market wage growth occurs at job changes. However they focus exclusively on white males with high attachment to the labor force. Loprest (1992) examines the difference in wage growth between men and women for job changers. She finds that men experience substantially higher wage gains at job changes. Our work advances these decompositions by focusing on lower skill male workers with weak attachment to the labor force. More importantly, constructing counterfactuals eliminating different types of turnover is more ambitious than decomposing wage growth. Topel and Ward's result is purely descriptive. The fact that 33% of wage growth occurs at job changes does not mean that wage growth would fall by 33% in the absence of turnover (and isn't interpreted in that way). In addition, these other papers do not account for the nonemployment spells accompanying turnover that are an important part of the cost of job turnover. More recent work by French, Mazumder, and Taber (2006) also decomposes components of wage growth by source. They use a sample of less educated workers in the Survey of Income and Program Participation. Although they focus mainly on changes over time, they do find a small net effect of turnover on wage growth.

The substantial literature estimating the return to tenure includes Topel (1986), Abraham and Farber (1987), Altonji and Shakoto (1987), Topel (1991), and Altonji and Williams (1998,2005). Most of these papers consider a regression model of the type (ignoring higher order terms and other covariates)

$$\log(w_{it}) = \theta_i + \beta_1 E_{it} + \beta_2 T_{it} + \eta_{ij_{it}} + \varepsilon_{ij_{it}t}, \tag{1}$$

where the individual specific effect θ_i is constant for a person across time and jobs, j_{it} is the job occupied by individual i at time t, the match specific component η_{ij} is assumed to be fixed over the length of the match, T_{it} is the level of tenure at job j_{it} at time t, and E_{it} is experience. The final error term represents measurement error or other unrelated variation in wages. The primary goal of this literature is to estimate β_2, the return to tenure. Estimates of β_2 vary considerably across papers. We use this framework in our work, but estimating β_2 is not essential for our approach. From a purely academic sense β_2 is an interesting parameter because it allows one to distinguish between theories of wage determination. However, from a program evaluation framework, β_2 can be interpreted as the difference between wages in the state of the world in which a worker stays on the same job versus the counterfactual in which he starts a new job with exactly the same job match component $(\eta_{ij_{it}})$. This is not a very interesting treatment. By contrast, our approach is to compare the returns to staying on the same job to the job match component that an individual actually receives when changing jobs. This essentially conditions on the selection bias rather than eliminating it. Even though this comparison is (arguably) easier to construct than estimation of β_2, we believe that this type of counterfactual is more useful for understanding the importance of turnover.

A few relatively recent papers have attempted to measure the effect of turnover on young workers. Gardecki and Neumark (1998) regress wages in the NLSY79 on the standard set of covariates and factors proxying the early labor market experience. They find little relationship between job stability of workers and their wages later in life. Neumark (2002) extends this work using early-in-life local labor market conditions as an instrument for job stability. The IV results point towards positive effects of job stability. Light and McGarry (1998) also estimates the effects of early career job mobility on wages. They use a panel data/instrumental variable approach and find that job mobility is negatively associated with wages.

While these papers are interesting, it is difficult to interpret the estimates. Instrumental Variables estimators have the cleanest interpretation when the coefficient on the endogenous variable is nonstochastic. However, when the endogenous variable is job turnover this is an implausible assumption. To see why, consider the following counter-factual using the wage specification (1). Let j_{it}^* represents the (typically counterfactual) job to which individual i would switch if he switched jobs at time t. The effect of a job change on the wage at time t for individual i is

$$\begin{aligned} &E(\log(w_{it}) \mid j_{it} = j_{it}^*) - E(\log(w_{it}) \mid j_{it} = j_{it-1}) \\ &= \eta_{ij_{it}^*t} - \eta_{ij_{it-1}t} - \beta_2 \left(T_{ij_{it-1}t-1} + 1\right). \end{aligned}$$

The assumption that this return is fixed is extremely strong. In fact, it is a central feature of job matching models that this difference is not constant. It is sometimes large and positive and other times very small and negative. Angrist, Graddy, and Imbens (2000) provide an interpretation of the IV estimator in a random coefficient case showing that the parameter can be written as a weighted average of this parameter across individuals. However, without a good understanding of precisely what those weights mean, it is difficult to know how to interpret the estimates. Rather than try to estimate a single parameter of the effect of turnover, we estimate the full distribution which allows us to simulate the counterfactuals of interest.

Our work also builds on the substantial literature on structural models of wage growth using search/matching models. Few of these models estimate the job mobility process, but instead focus on the search process to the first job. Exceptions are Miller (1984), Flinn (1986), Wolpin (1992), and Barlevy (2008). However our approach is quite different in that we do not attempt to estimate all of the structure of the model but instead focus on the aspects that are crucial for estimating the counterfactuals.

3 General Model and Counterfactuals

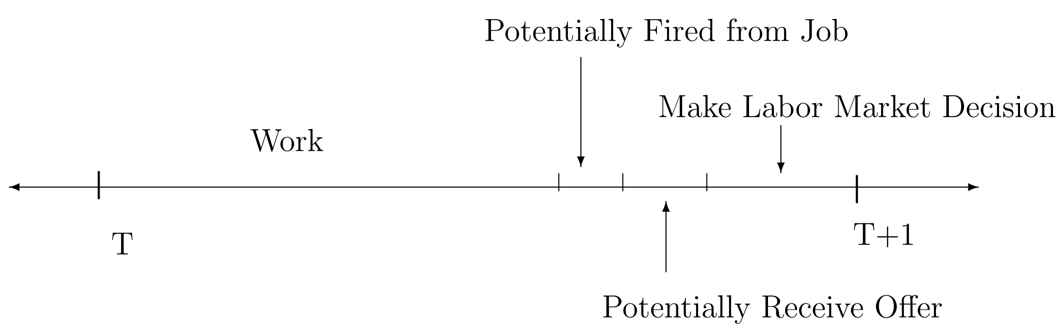
The state variables at any point in time for a worker i are $s_{it}, X_i, E_{it}, T_{it}, \eta_{ij_{it}}$, and θ_i where s_{it} is employment status at the beginning of time period t, X_i are observable demographic variables, $\theta_i = (\theta_i^w, \theta_i^u)$ are unobservable variables, E_{it} is total labor force experience, T_{it} is the amount of tenure at time t, and η_{ij} is the match component for worker i with firm j. Decisions at any time period t depends only upon the state variables and a vector of serially independent random variables.

We use somewhat nonstandard notation by letting j_{it} denote the job that individual i held at time t. We assume that firms post wages of the form

$$\log(w_{it}) = X_i'\gamma + \beta_1 E_{it} + \beta_2 T_{it} + \beta_3 E_{it}^2 + \beta_4 T_{it}^2 + \theta_i^w + \eta_{ij_{it}} + \varepsilon_{it} \tag{2}$$

where ε_{it} is a transitory error term. We assume that ε_{it} does not affect choices and do not model it explicitly. It can either be interpreted as measurement error or as a variable that is not realized until after labor supply decisions have been made.

First consider the case in which an individual is employed at the beginning of a period, the timing of the model is as follows:



A worker begins the period by working.[3] At the end of this sub-period the firm informs him whether or not he will be laid off. During the next subperiod he may receive an offer

[3] We imagine that the work subperiod takes the vast majority of the time.

from another firm.[4] The offer will contain a new value of the match component $\eta_{ij.}$ In the final sub-period the agent chooses among his options. As long as he was not laid off, he can choose to stay at the current firm. If he received an alternative offer, he can choose to accept it. If he was retained at his job, he has the option to quit to nonemployment.

At this point, a typical way to proceed would be to use a Bellman equation to define the value function explicitly. As we mentioned in the introduction we do not want to do this since it is unclear how to specify a value function for "floundering." Rather than estimating preferences per se, we estimate the behavior of agents with respect to the state variables and use this information in the simulations. Although the model is quite flexible, we do impose some structure. Most importantly we restrict the state space to be $(s_{it}, X_i, E_{it}, T_{it}, \eta_{ij}, \theta_i)$. We put no restriction on the relationship between these state variables and the way that decisions are made. However, when we perform the simulations later in the paper we make use of this structure in that we do not allow agents to use additional information to make decisions. For example, workers do not have additional information about their future job options. In the appendix (section A.2) we present a specific example of a model of foundering that fits this framework.

Our goal is to estimate the parameters of the wage equation as well as the parameters of the choice probabilities. As mentioned above, s_{it} denotes individual $i's$ working status at the beginning of time t where

$$s_{it} = \begin{cases} 0 & \text{individual } i \text{ works at time } t \\ 1 & \text{individual is not employed and entered this state through layoff} \\ 2 & \text{individual is not employed and entered this state through quit} \end{cases}.$$

It may be reasonable not to distinguish between $s_{it} = 1$ and $s_{it} = 2$ but we prefer to use a more general specification and let the data tell us whether this is a reasonable assumption.

[4] This can occur whether he was laid off or not although the arrival probability is different in the two cases.

Let ℓ_{it} indicate the labor market decision at the end of period t with

$$\ell_{it} = \begin{cases} 0 & \text{individual } i \text{ remains at same job} \\ 1 & \text{individual } i \text{ is laid off} \\ 2 & \text{individual } i \text{ voluntarily switches to new job} \\ 3 & \text{individual } i \text{ quits to nonemployment} \end{cases}.$$

It is important to note that the outcome $\ell_{it} = 1$ is slightly different than the others. For the other outcomes, knowledge of ℓ_{it} is sufficient to determine the evolution of the state variables in period $t+1$. However, when a worker is laid off he may immediately receive a new job and thus he may or may not be working in the following period.

To simplify the notation, write the state variables (other than s_{it}) for a worker as

$$I_{it}^e = (\eta_{ij_{it}}, E_{it}, T_{it}, X_i, \theta_i^w)$$

and the state variables for a worker without a job as

$$I_{it}^n = (E_{it}, X_i, \theta_i^u).$$

We then estimate the following conditional probabilities

$$\Pr(\ell_{it} = l \mid s_{it} = 0, I_{it}^e) \tag{3}$$

$$\Pr(s_{it+1} = 0 \mid s_{it} = 0, I_{it}^e, \ell_{it} = 1) \tag{4}$$

$$\Pr(s_{it+1} = 0 \mid s_{it} = 0, I_{it}^n) \tag{5}$$

for $l = 0, 1, 2, 3$.

The other main difference between our approach and a standard structural approach is that we estimate the <u>accepted</u> wage distribution rather than the wage offer distribution. To motivate this approach consider a very simple reservation wage labor model in which all workers face the same wage offer distribution and have the same reservation wage, w_r. Figure 1 presents the wage offer distribution and the reservation wage w_r. Given data on individuals with this reservation wage, one can observe the fraction of people who work and one can identify the distribution of accepted wages. In the figure, this corresponds to the distribution to the right of the solid line. However, we know nothing about the shape of the distribution of the wage offer distribution to the left of the reservation wage (other than it must integrate to the probability of not working).

Now consider a policy that leads to a lowering of the reservation wage w_r to the dashed line to the left of the solid line. Given that the distribution is not identified, we can say virtually nothing about the effect of the reservation wage change on either the employment fraction or the wage distribution. By contrast consider a shift in the reservation wage to the dashed line to the right. Since the accepted wage distribution is identified, for this case we can identify precisely how the fraction working changes as well as the change in the wage distribution. This simple example illustrates that some counterfactuals are identified from the accepted wage distribution and others are not.

Our formulation is substantially more complicated than the reservation wage example, but the same conditions hold. The full offer distributions are not identified, but the accepted wage distribution can be. Some counterfactuals can be identified in this case while others cannot. In this paper we will estimate the accepted wage distribution rather than the offered wage distribution. We concentrate on the counterfactuals that can be estimated from the accepted wage distribution after showing that these are identified.

For individuals who switch jobs we also estimate the distribution of the accepted new match component, $\eta_{j_{it+1}}$, conditional on $\ell_{it} = 2$ and the relevant state variables I_{it}.

Experience and tenure evolve nonstochastically in the standard way. If an individual works, experience is augmented by 1. If a person stays at the same job, tenure is augmented by 1. If the worker switches to a new job his tenure is reset to zero. We also make a distinction between workers who enter nonemployment through layoff and those who enter through quits: when $\ell_{it} = 3$, $s_{it+1} = 3$.

The focus of this paper is on separation from jobs, as opposed to the search decision of unemployed workers. With this in mind, we model the behavior of nonemployed workers in a simple manner. We abstract from the worker's job acceptance choice and just focus on the transition from non-employment to employment. As mentioned above, if a worker is laid off, there is some probability that he will immediately find a job so that he is not necessarily nonemployed in the following period. we estimate this as

$$\Pr(s_{it+1} = 0 \mid \ell_{it} = 1, I_{it}^{n}). \tag{6}$$

If a worker finds a new job, η_{ij} is drawn from a new distribution and $s_{it+1} = 0$. We allow the distribution to depend on whether the individual was laid off or quit their previous job.[5] If he does not find a job he does not work the following period and $s_{it+1} = 1$.

Workers who begin a period nonemployed find a job with probability

$$\Pr(s_{it+1} = 0 \mid s_{it}, I_{it}^{n}). \tag{7}$$

where this is only well defined for $s_{it} \in \{1, 2\}$.

Since we are not interested in searching for the first job, our model begins at the first post-schooling period in which the individual works. Thus $s_{i0} = E_{i0} = T_{i0} = 0$.

Our goal is not to estimate preferences of workers explicitly, but rather to estimate behavior as a function of the state variables. We define the behavioral model explicitly as follows.

Definition 1 *The **Behavioral Model** is characterized by the following properties*

- *Wages are determined according to*

$$\log(w_{it}) = X_i'\gamma + \beta_1 E_{it} + \beta_2 T_{it} + \beta_3 E_{it}^2 + \beta_4 T_{it}^2 + \theta_i^w + \eta_{ij_{it}} + \varepsilon_{it}$$

- θ_i *is i.i.d. with distribution* F_θ

- ε_{it} *is i.i.d. with distribution* F_ε

[5] In a previous version of the paper (Gladden and Taber, 2006) we imposed that they are the same and the results of the simulations are virtually identical.

- *Initial state variables are determined according to*
  - $E_{i0} = 0$
  - $T_{i0} = 0$
  - $\eta_{j_i(0)}$ *is i.i.d.. with distribution* $F_{\eta 0}$ *(which has mean zero)*
- *For employed individuals* ($s_{it} = 0$), ℓ_{it} *evolves according to* $\Pr(\ell_{it} = l \mid s_{it} = 0, I_{it}^{e})$. *The state variables then evolve as follows:*

  $\ell_{it} = 0 :\ s_{it+1} = 0, j_{it+1} = j_{it}, X_{it+1} = X_{it} + 1, T_{it+1} = T_{it} + 1$

  $\ell_{it} = 1 :\ X_{it+1} = X_{it} + 1, T_{it+1} = 0$

  With probability $\Pr(s_{it+1} = 0 \mid \ell_{it} = 1, I_{it}^{n})$ $s_{it+1} = 0$, *and* $\eta_{j_{it+1}}$ *is distributed* $F_{\eta 1}$, *otherwise* $s_{it} = 1$.

  $\ell_{it} = 2 :\ s_{it+1} = 0, X_{it+1} = X_{it} + 1, T_{it+1} = 0$,*and* $\eta_{ij_{it+1}}$ *has conditional distribution* $F_{\eta'}\left(\cdot; I_{it}^{e}\right)$

  $\ell_{it} = 3 :\ s_{it+1} = 3, X_{it+1} = X_{it} + 1, T_{it+1} = 0$
- *For nonemployed individuals* ($s_{it} \in \{1, 3\}$), $X_{it+1} = X_{it}, T_{it+1} = 0$. *A job is found with probability determined by* $\Pr(s_{it+1} = 0 \mid s_{it}, I_{it}^{n})$ *in which case* $s_{it+1} = 0$ *and* $\eta_{j_{it+1}}$ *is distributed* $F_{\eta s_{it}}$. *Otherwise* $s_{it+1} = s_{it}$.

The appendix contains a sketch of a proof that the behavioral model is nonparametrically identified. The fact that the model can be identified is not surprising. The parameters of the wage equation are identified under standard fixed effect types of assumptions. The rest of the model is identified because we have a lengthy panel. Our model restricts this correlation structure because we have a finite number of state variables. The more restricted model is identified from the general correlation pattern. Our claim above was that we was identifying something akin to the "accepted wage distribution" rather than the "wage offer distribution." This can see in the expressions above in that the various versions of $F_{\eta 1}$ and $F_{\eta'}$ represent the wage distributions rather than wage offers. In principle, this limits the counterfactuals we can simulate, but many interesting simulations can still be performed.

This behavioral model can be justified as a "reduced form" in the formal sense that we can write down a structural model and write these parameters as known functions of the structural model. Presumably there are more than one such model, but in appendix A.2 we present one model that formalizes the idea of 'floundering.'

We next discuss how to use this behavioral model to understand the role of turnover in wage growth; do workers turnover "too much," and what is the cost of this behavior? Our counterfactuals are not meant to represent literal policies that could be implemented. Rather they are thought exercises that are useful for understanding the underlying processes. All of these counterfactuals are identified from the behavioral model described above and some additional assumptions. One is that we are looking at partial equilibrium. We can think of the "policies" as affecting a single person drawn at average from the population.[6] Thus, for example, the wage posting policy that firms use will not change in response to the counterfactuals.

We find the following metaphor useful in interpreting the simulations. For each counterfactual, imagine that there is a job counselor who has a lot of power over the turnover decisions of the worker. For instance, the counselor could force the worker not to quit a job by having the ability to severely punish him. To simplify the exposition we let $\widetilde{\Pr}$ denote turnover probabilities under the simulated counterfactual and while Pr denotes the probability identified from the behavioral model.

No Turnover: The first counterfactual eliminates turnover completely. Formally we impose that

$$\widetilde{\Pr}(\ell_{it} = 0 \mid s_{it} = 0, I^e_{it}) = 1.$$

In this case the path is deterministic after the initial draw of random variables (thinking

[6]Heckman, Lochner, and Taber (1998) point out the importance of general equilibrium effects on labor market policy evaluation, but general equilibrium is beyond the scope of this particular paper.

of ε as measurement error). All workers receive wages according to (2). In terms of the job counselor metaphor, imagine that the counselor has the power both to prevent the worker from quitting and to prevent the firm from firing the worker. This is clearly not realistic, but does a good job in summarizing the overall effect of turnover. All that is needed to identify this counterfactual is the parameters of the wage equation, so the counterfactual is identified because the behavioral model is identified.

No Voluntary Turnover: The second counterfactual eliminates all voluntary turnover. Imagine that the counselor can prevent the worker from quitting, but has no power over the firm. Thus the layoff probability remains unchanged:

$$\widetilde{\Pr}(\ell_{it} = 1 \mid s_{it} = 0, I^e_{it}) = \Pr(\ell_{it} = 1 \mid s_{it} = 0, I^e_{it}).$$

However, since workers never quit we have:

$$\begin{aligned}\widetilde{\Pr}(\ell_{it} = 2 \mid s_{it} = 0, I^e_{it}) &= 0\\ \widetilde{\Pr}(\ell_{it} = 3 \mid s_{it} = 0, I^e_{it}) &= 0.\end{aligned}$$

This leaves

$$\widetilde{\Pr}(\ell_{it} = 0 \mid s_{it} = 0, I^e_{it}) = 1 - \Pr(\ell_{it} = 1 \mid s_{it} = 0, I^e_{it}).$$

Since $\Pr(\ell_{it} \mid s_{it} = 0, I^e_{it})$ is identified as part of the behavioral model, this counterfactual is identified.

No Voluntary Quits to Nonemployment: The third counterfactual eliminates all voluntary quits to non-employment. This case is similar to the previous one in that the counselor will not allow workers to quit jobs to enter nonemployment and in that firms are allowed to fire workers. However, in this case the counselor does not restrict the worker's job-to-job transitions.[7] Under this counterfactual the probability of staying on the job is

[7] However, we simulate the model under the assumption that the worker makes job-to-job transitions in the same way that they do in the estimated model. That is, we do not allow for strategic behavior from the agents where they can get around the restriction on quits to nonemployment by accepting a job from which they will likely be fired.

the sum of the estimated probability of staying on the job plus the probability of quitting to nonemployment. Formally,

$$
\begin{aligned}
&\widetilde{\Pr}(\ell_{it} = 0 \mid s_{it} = 0, I_{it}^e) = \Pr(\ell_{it} = 0 \mid s_{it} = 0, I_{it}^e) + \Pr(\ell_{it} = 3 \mid s_{it} = 0, I_{it}^e) \\
&\widetilde{\Pr}(\ell_{it} = 1 \mid s_{it} = 0, I_{it}^e) = \Pr(\ell_{it} = 1 \mid s_{it} = 0, I_{it}^e) \\
&\widetilde{\Pr}(\ell_{it} = 2 \mid s_{it} = 0, I_{it}^e) = \Pr(\ell_{it} = 2 \mid s_{it} = 0, I_{it}^e) \\
&\widetilde{\Pr}(\ell_{it} = 3 \mid s_{it} = 0, I_{it}^e) = 0.
\end{aligned}
$$

Nothing else in the model changes (including $F_{\eta'}$). Again, this is identified from the behavioral model.

Optimal Investment: Our next counterfactual is more involved. In this case, the counselor's job is to make optimal choices to maximize the worker's present value of earnings. However, the counselor only observes job offers that the worker reveals to her. She does not observe job offers that the worker does not want to tell her about. Thus I do not let the counselor force the worker to quit a job that he would not have otherwise quit. This is consistent our goal of understanding the effects of *reducing* turnover.

Since the job counselor only observes offers that the worker chooses to reveal, the base case is altered in two ways: i) When an individual is employed, the counselor observes offers that the worker would have accepted in the base model. This occurs with probability $\Pr(\ell_{it} = 2 \mid s_{it} = 0, I_{it}^e)$ and the match component is distributed according to $F_{\eta'}\left(\eta_{ij_i(t)}, I_{it}^2\right)$. In the base case the individual would switch jobs. In this case the counselor has the option to prevent him from switching. If the wage loss is sufficiently large the worker will be forced to stay in his current job. ii) Quitting to nonemployment is handled similarly. When the worker wants to quit, the counselor can stop him. However, she can not force the worker to quit.

Formally, we assume that the counselor maximizes the expected present value of earnings:

$$
E\left[\sum_{t=1}^{T}\left(\frac{1}{1+r}\right)^{t-1} 1\left(s_{it}=0\right) w_{it}\right]
$$

subject to

$$\begin{aligned}
&\widetilde{\Pr}(\ell_{it}=1 \mid s_{it}=0, I_{it}^{e}) = \Pr(\ell_{it}=1 \mid s_{it}=0, I_{it}^{e}) \\
&\widetilde{\Pr}(\ell_{it}=3 \mid s_{it}=0, I_{it}^{e}) \leq \Pr(\ell_{it}=3 \mid s_{it}=0, I_{it}^{e}) \\
&\widetilde{\Pr}(\ell_{it}=2, h_{\ell} \leq \eta_{ij_i(t+1)} \leq h_h \mid s_{it}=0, I_{it}^{e}) \\
&\qquad \leq \Pr(\ell_{it}=2 \mid s_{it}=0, I_{it}^{e}) \left[F_{\eta'}\left(h_h; I_{it}^{e}\right) - F_{\eta'}\left(h_{\ell}; I_{it}^{e}\right)\right] \forall \left\{(h_{\ell}, h_h) \in \Re^2 : h_{\ell} < h_h\right\}.
\end{aligned}$$

The last inequality incorporates the idea that the counselor can observe the wage offer before advising the client. The counselor can prevent the worker from taking "bad offers" but can not increase the number of accepted "good offers." It is in this sense that this model is related to the labor supply model presented in Figure 1. Offers that are turned down in our model are analogous to offers turned down in the labor supply model. Since we don't get data on them, we can say very little.

Simulating this model requires first estimating a dynamic programming model to maximize the present value of earnings. We then simulate the optimal paths. Again all of the pieces needed to simulate this counterfactual are identified from the reduced form.

In constructing the "optimal" profile we implicitly make a couple of important assumptions. First, we assume that a worker's preferences are unaltered by the counterfactual even though he knows his future decisions will be changed. Taken literally this is an extremely strong assumption. However, when we interpret the counterfactual as a thought experiment rather than as an actual policy it seems less strong. The point of this counterfactual is to answer the following question: How much could we increase wage growth for low wage workers if we could selectively reduce the amount of voluntary turnover leaving everything else fixed?

The second assumption is less obvious, but is probably more important. By using the wage model (2) we assume that we know the counterfactual wage that a worker would earn if he stayed. This is quite strong. A worker may have knowledge that is not captured by the observed data. If a worker anticipates never receiving a further wage raise or being laid off from a firm, he may choose to quit now. This may lead to a wage loss in the short run, but wage gains in the long run. To the extent that this is important, we will be overstating the negative impact of voluntary turnover. This is a difficult but potentially important consideration for future work.

Rule of Thumb: Our next counterfactual again allows the counselor to restrict turnover to increase the present value of earnings. Assume the counselor has exactly the same information, power, and restrictions as in the previous counterfactual. However, the counselor is not sophisticated enough to solve a dynamic programming model. Instead she just follows two rules:

1. The worker never quits to nonemployment:

$$\widetilde{\Pr}(\ell_{it} = 3 \mid s_{it} = 0, I^e_{it}) = 0.$$

2. The worker leaves a job with match quality η to accept a new job with match quality η' if and only if $\eta' > \eta$. Formally this can be written as

$$\widetilde{\Pr}(\ell_{it} = 2, \eta_{ij_{it+1}} \leq \eta_{ij_{it}} \mid s_{it} = 0, I^e_{it}) = 0$$
$$\widetilde{\Pr}(\ell_{it} = 2, \eta_{ij_{it}} < \eta_{ij_{it+1}} \leq h_h \mid s_{it} = 0, I^e_{it}) = \Pr(\ell_{it} = 2 \mid s_{it} = 0, I^e_{it}) \left[F_{\eta'}(h_h; I^e_{it}) - F_{\eta'}(\eta_{ij_{it}}; I^e_{it})\right]$$

4 Data

The empirical analysis uses data from the National Longitudinal Survey of Youth 1979 (NLSY79). This is a panel data set begun in 1979 with youth aged 14 to 22. The survey was conducted annually until 1994 when it started to be conducted biennially. Respondents are questioned on a large range of topics, including schooling, wages, and work experience. We use the cross-sectional sample as well as the oversamples of blacks and Hispanics. Because we are interested in early career wage growth for less educated workers, and to avoid complications of leaving the workforce for college, we use only individuals with 12 or fewer years of education, and we use only the first 10 years after an individual leaves school. Also, to avoid modeling the effect of child bearing on women's labor market decisions, we use only males. This leaves us a sample of 2431 individuals.

In constructing our data set, we make heavy use of the work history files. These provide detailed information on job turnover and employment. In particular, the work history files include information on the week that an individual started and ended each job, hourly wages on up to 5 jobs for each individual in each year, and information on the reason the individual left each job. This data set is ideal for this study for two reasons. First, it is necessary that we observe workers during their early years in the labor market. Second, the goal is to understand the consequences of early labor market decisions on wages later in life so it is also essential that we have a long panel.

One issue that arises when computing job turnover is determining which job changes to count. Many people in the sample hold more than one job at a time, or leave a job but return to it later. For our purposes, we decided not to count jobs that are obviously second jobs - that is, jobs that begin after, and end within, the time frame that another job is held. It is reasonable to expect that these jobs would have little effect on wage growth, since they are not primary jobs. However, if a person leaves a job for four or more weeks, we count this as a job separation even if he eventually returns to it. The reasoning is that people often hold a job or search for another job while they are away from this employer, so returning to a job might represent job shopping that didn't work out. The preliminary results were not sensitive to whether or not these breaks in a single job are counted as separations. When two jobs overlap, we assume that the second job began after the first ended. This assumption does not change the amount of turnover; it simply changes the length of the second job. The reasoning is that we are interested in primary jobs, and the second job becomes the primary job when the individual leaves the first job. We expect that this assumption does not change our basic results, but it is a necessary assumption since our econometric approach does not account for multiple job holding.

Table 1 presents summary statistics for our sample of individuals. Since we use the over-sample of black and Hispanic workers, they are over-represented. Because of the education cutoff, the sample has low levels of education. We observe most individuals for the full ten years: the average amount of time is 9.7 years. There are two reasons we don't observe some individuals for the full 10 years: some individuals left school before the survey began in 1979, and some individuals leave the survey before they have been out of school 10 years. The typical person in our sample works about 7.8 of the first 10 years they are out of school, and holds about 6.2 jobs during these 10 years. The average worker is fired 2.1 times, quits to nonemployment 1.3 times, and experiences 1.9 job-to-job transitions. A notable feature here is the number of quits to nonemployment. While quitting to nonemployment is rarer than being fired or quitting to a new job, it is not at all uncommon.

One potential explanation for some of the quits to nonemployment is that individuals have to quit a job when they are sent to jail or prison. Unfortunately, the NLSY79 does not collect detailed information on spells of incarceration, so we cannot fully control for this issue. The only information that we have is whether an individual is incarcerated at the time of the yearly interview. To get a sense of what portion of quits to non-employment could be caused by spells of incarceration, we construct a yearly panel and define a variable that takes the value of 1 if, at the time of the interview, an individual is in a spell of non-employment spell caused by a quit. In 1.2% of the person-years in our sample the individual is incarcerated at the time of the interview. Individuals who are in jail at the time of the interview are much more likely to have quit to non-employment than other individuals in the sample: 55% of incarcerated men versus 5.5% of other men have quit to non-employment. However, incarcerated men account for only 11% of all quits to non-employment. This is an upper bound on the amount of voluntary quits to nonemployment caused by incarceration because in some cases, the quit to nonemployment could have preceded the arrest. Although ideally we would have the data to better control for this issue, given the relatively small portion of quits to non-employment that are in jail at the time of the interview we proceed with the data that we have.

The main empirical work in the paper focuses on wage growth within and between jobs. As a result, the unit of observation is not a person-year (as is typical in panel data work) but a job. Table 2 summarizes the demographic characteristics of job changers and the effects of job changes on wages. We look at three types of separations: employer initiated, job-to-job transitions, and job-to-nonemployment transitions. A worker typically experiences a decline in wages when he is fired and a large increase at a job-to-job transition. Once again, quits to nonemployment are not uncommon. When a worker quits to nonemployment he usually sees a wage increase at his next job. Workers who are white and who have more education, tenure, and experience are more likely to experience job-to-job transitions and less likely to experience involuntary and job-to-nonemployment transition.

Our behavioral model is written in discrete time and we need data in a form that is consistent with the model. The underlying data is weekly. In choosing the length of a period there is a trade-off between the computational cost of estimating the model versus the amount of information that is lost with too broad an aggregation. We choose 4 weeks as a period in our model meaning that there are 13 periods per year. Since search for the first job is not an important component of this model it is natural for period zero to correspond to the first period in which the individual worked after leaving school. We follow people for total of ten years or 130 4-week periods.

5 Estimation of Parameters of the Model.

In this section we discuss the estimation of the model. We estimate the model in three stages. First we estimate the parameters of the wage equation. Second we estimate the parameters determining turnover and the distribution of the error terms (other than $F_{\eta'}\left(\cdot; I_{it}^{e}\right)$). Finally we estimate the parameters governing $F_{\eta'}\left(\cdot; I_{it}^{e}\right)$. The reason for separating the second and third steps is described below.

First, consider the estimation of the wage equation. For job stayers between time t and

τ (i.e. individuals for which $j_{it} = j_{i\tau}$),

$$\log(w_{i\tau}) - \log(w_{it}) = \beta_1 (E_{i\tau} - E_{it}) + \beta_2 (T_{i\tau} - T_{it}) + \beta_3 \left(E_{i\tau}^2 - E_{it}^2\right) + \beta_4 \left(T_{i\tau}^2 - T_{it}^2\right) + \varepsilon_{i\tau} - \varepsilon_{it}.$$

A key issue in estimating this model is the well known problem that for job stayers $(E_{i\tau} - E_{it}) = (T_{i\tau} - T_{it})$ so we can not identify all of the parameters. However, we can use job stayers to get a consistent estimate of $(\beta_1 + \beta_2, \beta_3, \beta_4)$.

To separate β_1 from β_2 we make use of the implication of the model that immediately after a nonemployment spell, tenure on the previous job and experience are unrelated to the match η_{ij} that is drawn. While this may be a strong assumption in other contexts, it is important to keep in mind that identification of β_2 is not crucial for our counterfactual simulations. What matters most is the distribution of wage changes that occurs at job transitions. One could allow β_2 to be larger and estimate the mean value of η on the next job. The mean value of the wage on the next job will increase as β_2 gets larger, leaving the distribution of wage changes at job switches at a similar level. It is this distribution which is most important for the simulations.

To make our estimation approach more explicit let

$$\begin{aligned}
t_q^* &: \text{ first period working following first quit} \\
t_\ell^* &: \text{ first period working following first layoff} \\
t_1^* &: \text{ first period working following first quit or layoff} \\
t_2^* &: \text{ first period working following second quit or layoff}
\end{aligned}$$

Let μ_q and μ_ℓ be the average draw of η_{ij} following a nonemployment spell initiated from a quit or a layoff respectively. If we observed wages just before and after a nonemployment spell we could use the moment equations

$$\begin{aligned}
E\left[\log(w_{it_q^*}) - \log(w_{i0})\right] &= \beta_1 E\left(E_{it_q^*}\right) + \beta_3 E\left(E_{it_q^*}^2\right) + \mu_q \\
E\left[\log(w_{it_\ell^*}) - \log(w_{i0})\right] &= \beta_1 E\left(E_{it_\ell^*}\right) + \beta_3 E\left(E_{it_\ell^*}^2\right) + \mu_\ell \\
E\left[\log(w_{it_2^*}) - \log(w_{it_1^*})\right] &= \beta_1 E\left(E_{it^*}\right) + \beta_3 E\left(E_{it^*}^2\right) + \mu_{i2} - \mu_{i1}
\end{aligned}$$

to estimate the parameters $(\beta_1, \mu_q, \mu_\ell)$ where μ_{i1} and μ_{i2} are μ_q or μ_ℓ depending on whether the first and second spells correspond to a quit or a layoff.

In practice wages are observed at interview dates rather than at the start and end of jobs. Let τ_s^* represent the amount of tenure a worker has when we observe the wage for $s \in \{0, q, \ell, 1, 2\}$. Then one obtains the equations:

$$\begin{aligned}
&E\left[\log(w_{i\left(t_q^*+\tau_q^*\right)}) - \log(w_{i\left(0+\tau_0^*\right)})\right] \\
&\quad - (\beta_1+\beta_2)\left(\tau_q^* - \tau_0^*\right) - \beta_3 E\left(E_{i\left(t_q^*+\tau_q^*\right)} - E_{i\left(0+\tau_0^*\right)}\right)^2 - \beta_4\left(\tau_q^* - \tau_0^*\right)^2 \\
&\quad = \beta_1 E\left(E_{i\left(t_q^*+\tau_q^*\right)} - E_{i\left(0+\tau_0^*\right)} - \left(\tau_q^* - \tau_0^*\right)\right) + \mu_q \qquad (8) \\
&E\left[\log(w_{i\left(t_\ell^*+\tau_\ell^*\right)}) - \log(w_{i\left(0+\tau_0^*\right)})\right] \\
&\quad - (\beta_1+\beta_2)\left(\tau_\ell^* - \tau_0^*\right) - \beta_3 E\left(E_{i\left(t_\ell^*+\tau_\ell^*\right)} - E_{i\left(0+\tau_0^*\right)}\right)^2 - \beta_4\left(\tau_\ell^* - \tau_0^*\right)^2 \\
&\quad = \beta_1 E\left(E_{i\left(t_\ell^*+\tau_\ell^*\right)} - E_{i\left(0+\tau_0^*\right)} - \left(\tau_\ell^* - \tau_0^*\right)\right) + \mu_\ell \qquad (9) \\
&E\left[\log(w_{i\left(t_2^*+\tau_2^*\right)}) - \log(w_{i\left(t_1^*+\tau_1^*\right)})\right] \\
&\quad - (\beta_1+\beta_2)\left(\tau_2^* - \tau_1^*\right) - \beta_3 E\left(E_{i\left(t_2^*+\tau_2^*\right)} - E_{i\left(t_1^*+\tau_1^*\right)}\right)^2 - \beta_4\left(\tau_2^* - \tau_1^*\right)^2 \\
&\quad = \beta_1 E\left(E_{i\left(t_2^*+\tau_2^*\right)} - E_{i\left(t_1^*+\tau_1^*\right)} - \left(\tau_2^* - \tau_1^*\right)\right) + \mu_{i2} - \mu_{i1} \qquad (10)
\end{aligned}$$

we can obtain a consistent estimate of $(\beta_1, \mu_q, \mu_\ell)$ by using sample analogues of the expected values. This identification strategy is related to that used by Topel (1991). The first stage is identical. The second differs in that we only look across spells when a spell was terminated by a layoff or a quit to nonemployment. As in Topel's (1991) procedure, our estimate of tenure will tend to be biased downward to the extent that E_{it^*} is positively correlated with $\eta_{ij_{it^*}} - \eta_{ij_{i0}}$. For example, this would occur if workers engage in directed search and labor market experience is informative for jobs that are likely to be the best matches.

The estimates of this procedure are presented in Table 3, Panel 1. We find positive returns to experience and returns to tenure that are essentially zero.

Although we have shown that the behavioral model is nonparametrically identified, with a finite data set a completely nonparametric approach is infeasible. We therefore choose a flexible parametric model. There are five different aspects of the model that need to be estimated. The first four sets of these parameters were estimated at once using simulated maximum likelihood. For expositional purposes we focus on each of the pieces separately, but one should keep in mind that the parameters were estimated simultaneously.

First consider estimation of the other parameters of the wage equation (γ in equation 2). These estimated parameters are presented in the first panel of Table 3. Not surprisingly, black workers and less educated workers are paid less than others. The Hispanic coefficient is positive, but statistically insignificant.

Second, consider the distribution of the error terms. We assume that

$$\begin{aligned} F_\theta &= \Phi(\cdot;\mu_\theta,\sigma_\theta^2), F_\varepsilon = \Phi(\cdot;0,\sigma_\varepsilon^2), F_{\eta 0} = \Phi(\cdot;0,\sigma_{\eta 0}^2), \\ F_{\eta 1} &= \Phi(\mu_\ell;0,\sigma_{\eta 1}^2), F_{\eta 3} = \Phi(\mu_q;0,\sigma_{\eta 3}^2) \end{aligned}$$

where $\Phi(\cdot;\mu,\sigma^2)$ represents the c.d.f. of a normal random variable with mean μ and variance σ^2.

The parameters of these distribution functions are presented in the bottom half of Table 3, Panel 1. We see no surprises in any of these parameters.

The next set of parameters governs the transition probabilities $\Pr(\ell_{it} = l \mid s_{it} = 0, \eta_{ij_i(t)}, E_{it}, T_{it}, X_i, \theta_i^w)$. Let $Z_{it} = \begin{bmatrix} E_{it} & T_{it} & E_{it}^2 & T_{it}^2 \end{bmatrix}'$. We model these using the following multinomial logit approximation,

$$\begin{aligned} &\Pr(\ell_{it} = l \mid s_{it} = 0, \eta_{ij_{it}}, E_{it}, T_{it}, X_i, \theta_i^w) \\ &\quad = \frac{\exp(X_{it}'\lambda_x^\ell + \lambda_\theta^\ell\theta_i^w + \lambda_\eta^\ell\eta_{ij_{it-1}} + Z_{it}'\lambda_z^\ell)}{1 + \sum_{s=1}^3 \exp(X_{it}'\lambda_x^\ell + \lambda_\theta^\ell\theta_i^w + \lambda_\eta^\ell\eta_{ij_{it-1}} + Z_{it}'\lambda_z^\ell)} \end{aligned}$$

for $\ell = 1, 2, 3.$[8]

[8]Note that this is somewhat more general than it may appear. One could in principle estimate this specification nonparametrically by allowing the complexity of the linear aspect of the model to increase with sample size (e.g. by adding higher order polynomial terms). This would be completely nonparametric

The parameters of this model are presented in Table 3, Panel 2. The results seem quite reasonable and are consistent with Table 2. Blacks and less educated workers are more likely to be fired and more likely to quit to non-employment. Blacks and Hispanics are less likely to make job-to-job transitions. Tenure reduces turnover of all types, while experience reduces quits to non-employment. As expected, the individual random effect θ^w is negatively associated with being fired or quitting to non-employment. Again quite intuitively, for all three transitions lower values of the match quality (η) make leaving the job more likely.

Similarly, we model the probability of finding a job using a logit model. We estimate the probability of finding a new job for the three different possible transitions: finding a job immediately, finding a job from nonemployment after a quit, and finding a job from nonemployment after being fired. We allow for another random effect θ_i^u which is related to an individual's turnover probability. θ_i^u is restricted to be normal and correlated with θ_i^w so that if can be represented as:

$$\theta_i^u = \alpha_1 \theta_i^w + \alpha_2 \theta_i^{u*}$$

where θ_i^{u*} is standard normal and uncorrelated with θ_i^w. Rather than estimate the components in multiple sets we allow for another random effect θ_i^{u*} and allow both θ_i^w and θ_i^{u*} to enter the logit equations. Our model is somewhat more general than this since we allow the α weights to vary with the type of turnover.

The estimates from these logits are presented in Table 3, Panel 3. The results are consistent with conventional wisdom. White workers, more experienced workers, and more educated workers are more likely to find a job immediately, and more likely to find a job after being fired or after quitting to non-employment.

so that the fact that we am using a multinomial logit approximation (e.g. the independence from irrelevant alternatives assumption) does not restrict the model. However, in the data that we have, with four different choices we don't have enough observations to credibly estimate a very general form so we keep the index linear.

The four components of the model described above were estimated simultaneously. The next component was estimated separately for reasons described below.

The most complicated aspect of our estimation procedure is estimating the distribution of the new match draw conditional on switching $F_{\eta'}\left(\eta_{ij_i(t)}; E_{it}, T_{it}, X_i, \theta_i^w\right)$. We want to use a flexible specification because in the dynamic programming simulation we will allow workers to choose optimal cutoffs based on the full state vector. However, the specification needs to be computationally feasible. To see how we parameterize this function, let P be a uniform (0,1) random variable. For any distribution F, there exists a function g such that $g(P)$ has distribution F. For example, if the c.d.f. is invertible then $g = F^{-1}$ is one such function, but there are many others. Our goal is to estimate g rather than to estimate $F_{\eta'}$ directly. This formulation is useful because it makes our simulation straight forward. To simulate the draws of η' we simply draw a uniform random variable and transform it using g.[9] In particular we assume that g is a k-degree polynomial,

$$g\left(P; E_{it}, T_{it}, X_i, \theta_i^w\right) = \sum_{\ell=0}^{k} a_\ell\left(E_{it}, T_{it}, X_i, \theta_i^w\right) \log\left(\frac{P}{1-P}\right)^\ell. \tag{11}$$

In the expression above, if $k = 1$, $a_0 = 0$, and $a_1 = 1$ then η' is logistic.[10] This approach simplifies the computation of the model significantly for two reasons. First, estimation by simulated maximum likelihood becomes straightforward since the simulated distribution of the random variable will be continuous in the parameters a_ℓ. This guarantees that once we draw a set of uniform random variables the simulated likelihood function will also be smooth in its parameters. This would not be the case if we parameterized the model using an alternative such as a distribution of normals in which the likelihood function would jump when the parameters determining which node occurs jumps. As is the usual case with simulated maximum likelihood we need to assume that the number of simulations is large in order to have a consistent estimate.

[9] In fact, numerical random variable simulators typically operate in precisely this manner.

[10] Note that in the parameterization we have chosen we am not imposing that g be monotonic as it would be if $g = F^{-1}$. We experimented a bit with imposing monotonicity but found no advantage in doing this.

Second, after estimating this object, performing simulations is straight forward. We can construct simulations using our estimates by first drawing a uniform and then applying this function.

In practice we implement this approach using a third degree polynomial ($k = 3$) with the following restrictions on the values of $a_\ell(E_{it}, T_{it}, X_i, \theta_i^w)$ defined in equation (11). For the intercept $k = 0$, I take

$$a_0(E_{it}, T_{it}, X_i, \theta_i^w) = \rho_{00} + \rho_{\theta 0}\theta + \rho_{\eta 0}\eta + \rho_{\delta 0}X'\delta_0 + \rho_{\theta\eta}\theta\eta + \rho_{\theta\delta}\theta X'\delta_0 + \rho_{\eta\delta}\eta X'\delta_0. \quad (12)$$

For $k > 0$ we restrict

$$a_\ell(E_{it}, T_{it}, X_i, \theta_i^w) = \rho_{0\ell} + \rho_{\theta\ell}\theta + \rho_{\eta\ell}\eta + \rho_{\delta\ell}X'\delta_1. \quad (13)$$

We estimate both the ρ parameters and the δ parameters and present them in Table 3, Panel 4. Given the flexibility of the approximation, the parameters are hard to interpret directly, but they are important in generating the estimated distribution of job switches.

As a practical matter estimation of these parameters is quite difficult. The parameterization is very flexible. This allows the model to fit the data well, but it also means that more than one parameterization of the model might fit the data well. Thus one might expect local optima to be a severe problem and in fact it was. This is why we separate stages 2 and 3 in estimation. We estimate the other parameters of the model using maximum likelihood ignoring changes in the value of η after the first period. For this part of the problem, local optima are not a problem, but it takes a relatively long time to compute the likelihood function. In stage 3 we estimate the parameters of η taking the other parameters as given and use maximum likelihood based upon the distribution of wage gains. Since we focus on only part of the data (job-to-job switches) we can evaluate the likelihood function much more quickly.[11] This speeds up the estimation and facilitates finding the global optima.

[11] It is not only job-to-job switches that are relevant here since I need to use the model to control for selection. However, we can ignore searching off the job which is not relevant for identification of these parameters.

However, in practice finding the global optima proved very difficult. We experimented with a Simulated Annealing algorithm which is designed to find global optima, but we had little success. We found a less sophisticated method that worked much better. We randomly drew starting parameters for the model and estimated the likelihood function 4000 times.[12] Unfortunately the maximum of this procedure was found only one time in 4000 leaving little confidence it is the true global maximum. However, we simulated the model for the three highest values and found simulations results that are almost identical. Thus the different local optima seem to reflect the fact that there are many different parameterizations that approximate the key features of the data well so that finding the true global maximum is not essential.

To see how well we approximate the data, in Figure 2 we plot the fit of the model in various dimensions. The first panel presents the predicted wages of workers. The next few panels present the transition probabilities: nonemployment-to-employment, employment-to-nonemployment, and job-to-job. These all fit the actual data very closely (other than in the first few periods). The fit did not come as a direct implication of the first order conditions from the model because we plot the results by age, but the model conditions on experience and tenure and there is a lot of selection so it is not "rigged" to fit well.

6 Simulation Results

We generate two different simulation samples. In the first we draw 4000 different values of $(X_i, \theta_i^w, \theta_i^u)$ and then simulate the model 100 times for each. The results of these 400,000 simulations are presented in Figure 3a and 3b. Each figure presents four simulations as discussed in Section 3 above: a base case, a case eliminating all turnover, a case eliminating all voluntary turnover, and a case elimination all voluntary quits to non-employment.

The simulated wage profiles appear in Figure 3a. In order to reduce the importance of selection bias in explaining the differences, we present a "wage" for all individuals. For

[12] This took about two months running simultaneously on four processors.

men who are not working, the wage is set to be the last wage they received when they did work. This figure clearly illustrates that turnover is an important component of wage growth. When either all turnover or all voluntary job-to-job turnover is eliminated, wage growth is substantially reduced. Specifically wages are 7.3% lower if voluntary turnover is eliminated. By contrast, quitting to nonemployment does not seem to be productive. If voluntary quits to non-employment are eliminated, wage growth is steeper than in the base case.

The simulated earnings profile in Figure 3b tells a somewhat different story. We treat earnings as zero for nonworkers. Earnings in the counterfactuals differ from earnings in the base case for two reasons. First, in all three counterfactuals earnings increase because reducing turnover reduces time not working. Second, reducing turnover increases wage growth (and therefore earnings) when only quits-to-non-employment are eliminated, but decreases wage growth and earnings in the other two counterfactuals. In the early years in the labor market, the first effect is clearly dominant. All three counterfactuals predict higher annual earnings for the first 7 years in the labor market compared to the base case.

However, in later years earnings are affected by the reduction of wages that comes from eliminating too must turnover. By year 7, predicted annual earnings are higher in the base case than in the counter-factual eliminating all voluntary turnover. This is because eliminating voluntary turnover leads to lower wage growth, reducing earnings. Eliminating all turnover increases earnings in all of the first 10 years in the labor market. However, by the 10th year wages are higher in the base case than for either the no turnover or the no voluntary turnover case, so future earnings are also likely to be higher in the base case.

At the bottom of figure 3b, we present the present value of earnings over the first 10 years in the labor market for the four counterfactuals using a 4% annual interest rate. The base case leads to the lowest level of present value of earnings. Eliminating voluntary turnover increases earnings over the base case primarily because of the lost income associated with quits to nonemployment. It should be pointed out that this is only earnings from the first ten years. In the last period, wages are higher in the base case than for either the no turnover or the no voluntary turnover case, so future earnings are likely to be higher (assuming that the fraction of the sample that works continues to increase with experience). Another key issue is that this calculation essentially values leisure at zero. Another extreme would be to assume that the value of leisure can be approximated by the wage. With this in mind we also calculate the present value of wages for each of the four cases where as above, the wage of a non-worker is taken to be the last wage he received when he did work. These results are presented in the table at the bottom of the figure. Clearly they give a somewhat different story. If the value of leisure is somewhere between zero and the wage, then the "true" value is somewhere between the two. Nevertheless, the present value of wage results clearly demonstrate that overall, turnover is detrimental for wage growth.

Our next set of simulations calculates the optimal profile. In this case we solve the dynamic programming problem to uncover the optimal turnover strategy in all 130 periods for every value of $(X_i, \theta_i^w, \theta_i^u)$. This turns out to be very computationally costly, so rather than simulate for 4000 values of $(X_i, \theta_i^w, \theta_i^u)$ we only solve for 80. However, once the dynamic programming problem has been solved, it is not computationally costly to simulate the model. This allows us to simulate the model many many times (100,000) for each of these 80 values. The results of these simulations are presented in Figures 3a and 3b. we simulate 3 different models: (i) the base case[13]; (ii) An optimal strategy where workers change jobs to maximize the present value of earnings in the first 10 years with the restrictions discussed in section 3; and (iii) the "rule of thumb" described in section 3 where workers *a*) do not quit to nonemployment, *b*) change jobs only if it improves their match.

Figure 4a presents the simulated wage profiles. There is very little difference between the optimal strategy and the rule of thumb strategy. However, both of these strategies dominate the base case. Wages are higher in all periods, and are about 10% higher after workers have been in the labor market for 10 years.

[13]The simulation from base case is slightly different than before because the draws of $(X_i, \theta_i^w, \theta_i^u)$ in the simulations are a bit different.

Figure 4b presents simulated annual earnings. Again, there is little difference between the optimal strategy and the rule of thumb strategy. Both lead to a higher annual earnings in all years. After workers have been in the labor force for 10 years the optimal case predicts annual earnings of almost \$25,000 compared to annual earnings of about \$20,000 in the base case. Since wages are higher in the optimal case than in the base case, future earnings are also likely to be higher. The bottom panel of Figure 4b shows the present value of earnings over the first 10 years in the labor market for the three simulated cases using a 4% annual interest rate. Again, the simulations restricting turnover increase earnings compared to the base case: the present value of earnings are about 20% higher than under the "optimal" strategy. However, the present value of wages are only about 4% higher.[14]

7 Conclusions

The goal of this paper is to measure the costs of turnover for young male low skilled workers. To what extent does turnover represent "floundering" through jobs and to what extent does it resemble efficient human capital investment? The approach of the paper is to first estimate a behavioral model of turnover and wage growth, and to then use the results of the model to simulate various counterfactuals in which behavior is modified.

The first type of counterfactual quantifies the effect of eliminating various types of turnover on wage growth. Our results indicate that the overall effect of turnover on wage growth is positive, but workers do not simply maximize the present value of earnings. Completely eliminating all turnover substantially decreases wage growth but increases the present value of earnings over the first 10 years in the labor market due to the additional employment. Eliminating only voluntary turnover leads to lower wage growth but a somewhat higher present value of earnings. Finally, eliminating only voluntary quits to non-employment yields both higher wage growth and a higher present value of earnings.

The second type of counterfactual we construct is an "optimal" profile. Given the

[14]Present value of wages are calculated in the analogous way to that of Figure 3b.

observed distribution of wage gains in the data, we solve the dynamic programming problem to uncover the pattern of turnover that maximizes the present value of income. This exercise restricts turnover to only include productive switches. This yields some important lessons. We do not know the objective function of the workers, so we cannot say whether they turnover "too much." However, we do find evidence that workers are not purely maximizing the present value of earnings, and that reducing turnover would in fact increase both wages and earnings over the first 10 years. The distinction between what they actually do and what they "could do" is about 10% when measured by wage growth after 10 years and about 20% when measured as the present value of earnings during the first ten years in the labor force.

The fundamental question is whether workers appear to be 'floundering' or whether they are simply maximizing the present value of earnings. Our bottom line is that turnover behavior is somewhere in between. Eliminating all voluntary turnover would lead to wages that are 7.3% lower after ten years so workers are clearly not just 'floundering.' However, if you could force workers to decrease their voluntary turnover in an optimal way, their resulting wages would be 10% higher after ten years so they are not purely maximizing their income either. Thus both effects are sizable. Once again, it is important to point out that the deviation from the "optimal" case certainly does not mean that workers are not rational as one could surely find a set of preferences to justify the decisions. However, it does leave some scope that policy intervention could be justified. At the same time it is important to keep in mind that on net, average turnover is productive, so interfering could do more harm than good.

References

Abraham, Katherine and Henry Farber, 1987, "Job Duration, Seniority, and Earnings," *American Economic Review*, 278-297.

Altonji, Joseph, and R. Shakato, 1987, "Do Wages Rise with Job Seniority?" Review of Economic Studies, 437-459.

Altonji, Joseph and Nicolas Williams, 2005, "Do Wages Rise With Job Seniority? A Reassessment.", *Industrial and Labor Relations Review*, 58, Issue 3, Article 4.

Altonji, Joseph, and Nicolas Williams, 1998, "The Effect of Labor Market Experience, Job Seniority, and Job Mobility on Wage Growth," Research in Economics, Vol. 17, 223-276.

Angrist, Joshua, Kathryn Graddy, and Guido Imbens, 2000, "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish," *Review of Economic Studies*, 67, 499-527.

Antel, John, 1991, "The Wage Effects of Voluntary Labor Mobility with and without Intervening nonemployment," *Industrial and Labor Relations Review,* 44, 299-306.

Barlevy, Gadi, 2008, "Identification of Search Models using Record Statistics," forthcoming emphReview of Economic Studies.

Cameron, Stephen, and Christopher Taber, 1998, "Evaluation and Identification of Semiparametric Maximum Likelihood Models of Dynamic Discrete Choice," unpublished manuscript, University of Wisconsin-Madison.

Devine, Theresa, and Nicholas Kiefer, 1991, *Empirical Labor Economics,* Oxford University Press.

Farber, Henry, 1994, "The Analysis of Interfirm Worker Mobility," *Journal of Labor Economics,* 12, 554-593.

French, Eric, Bhashkar Mazumder, and Christopher Taber, 2006, "The Changing Pattern of Wage Growth for Low Skilled Workers ," unpublished manuscript.

Flinn, Christopher, 1986, "Wages and Job Mobility of Young Workers," *Journal of Political Economy,* 94, S88-S110.

Flinn, Christopher, and James Heckman, 1982, "New Methods for Analyzing Structural Models of Labor Force Dynamics," Journal of Econometrics, 115-168.

Gardecki, Rosella and David Neumark, 1998, "Order from Chaos? The Effects of Early Labor Marked Experiences on Adult Labor Market Outcomes," *Industrial and Labor Relations Review,* 51, 299-322.

Gladden, Tricia, and Christopher Taber, 2000, "Wage Progression Among Low Skilled Workers," *Finding Jobs: Work and Welfare Reform*, Card and Blank eds., 160-192.

Gladden, Tricia, and Christopher Taber, 2006, "Turnover and Wage Growth in the Transition from School to Work," unpublished manuscript.

Heckman, James, 1993, "Assessing Clinton's Program on Job Training, Workfare, and Education in the Workplace," NBER Working Paper No. 4428.

Heckman, James, and Burton Singer, 1984, "A Method for Minimizing the Impact of Distributional Assumptions in Economic Models for Duration Data," *Econometrica*, 271- 320.

Ichimura, Hidehiko, and Christopher Taber, 2001, "Direct Estimation of Policy Effects," unpublished Manuscript.

Ichimura, Hidehiko, and Christopher Taber, 2002,"Semiparametric Reduced Form Estimation of Tuition Subsidies," with H. Ichimura, *American Economic Review, Papers and Proceedings*, May.

Keane, Michael, and Kenneth Wolpin, 1994, "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpretation: Monte Carlo Evidence," *Review of Economics and Statistics,* 76, 648-672.

Klerman, Jacob, and Lynn Karoly, 1994, "Young Men and the Transition to Stable Employment," *Monthly Labor Review,* 31-48.

Kotlarski, Ignacy, 1967, "On Characterizing the Gamma and Normal Distributions, Pacific Journal of Mathematics, vol 20, 69-76.

Laibson, David, 1997, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, Vol. 112, No. 2, 443-477.

Light, Audrey and Kathleen McGarry, 1998, "Job Change Patterns and the Wages of Young Men," *Review of Economics and Statistics,* 276-286.

Loprest, Pamela, 1992, "Gender Differences in Wage Growth and Job Mobility," *American Economic Review,* 526-532.

Miller, Robert, 1984, "Job Matching and Occupational Choice," *Journal of Political Economy*, 1086-1120.

Mincer, Jacob, 1986, "Wage Changes in Job Changes," *Research in Labor Economics,* vol 8 Ehrenberg ed., JAI Greenwich CT.

Mincer, Jacob, and Boyan Jovanovic, 1981, "Labor Mobility and Wages," in Sherwin Rosen (ed), *Studies in Labor Markets,* Chicago University Press.

Neumark, David 2002, "Youth Labor Markets in the US: Shopping Around Vs. Staying Put," *Review of Economics and Statistics*, Vol. 84, No. 3, 462-484.

Prakasa Rao, B.L.S., 1992, *Identifiability in Stochastic Models*, Academic Press.

Stern, David, Neal Finkelstein, James Stone, John, and Carolyn Dornsife, 1994, "Research on School-to-Work Transition Programs in the United States," National Center for Research in Vocational Education Report.

Topel, Robert, 1986, "Job Mobility, Search, and Earnings Growth: A Reinterpretation of Human Capital Earnings Functions," *Research in Labor Economics,* vol 8 Ehrenberg ed., JAI Greenwich CT.

Topel, Robert, 1991, "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority," *Journal of Political Economy*, 145-176.

Topel, Robert and Michael Ward, 1992, "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics,* 108, 439-479.

Wolpin, Kenneth, 1992, "The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth," *Journal of Political Economy,* 100, 535-560.

Wolpin, Kenneth, 1995, "Empirical Methods for the Study of Labor Force Dynamics," Harwood Academic Publishers.

Appendix

A.1 Sketch of Proof of Identification

In this appendix, we provide an informal discussion on the identification of the parameters $\gamma, \beta, F_\theta, F_{\eta 0}, F_{\eta 1}, F_{\eta 3,}$ and F_ε from equation (2)

$$\log(w_{it}) = X_i'\gamma + \beta_1 E_{it} + \beta_2 T_{it} + \beta_3 E_{it}^2 + \beta_4 T_{it}^2 + \theta_i^w + \eta_{ij_i(t)} + \varepsilon_{it}$$

as well as the relevant conditional distribution of ℓ_{it} and the job finding probabilities. Throughout, we assume that we observe people for a long panel.[15] We also assume that for each individual i, θ_i is independent of X_i.

First consider identification of $\gamma, \beta, F_\theta, F_{\eta 0}, F_{\eta 1}, F_{\eta 3,}$ and F_ε.

(i) Identifying γ is straight forward. Since it is uncorrelated with the error term, γ is identified from a regression of the first period wage on X_i.

(ii) $(\beta_1 + \beta_2), \beta_3,$ and β_4 are identified from fixed effects on the wage equation (2). Since we have multiple observations on individuals within a job spell these parameters are identified from the regression:

$$E\left(\log(w_{i\tau}) - \log(w_{it}) \mid j_{it} = j_{i\tau}\right) = (\beta_1 + \beta_2)(\tau - t) + \beta_3 \left(E_{i\tau}^2 - E_{it}^2\right) + \beta_4 \left(T_{i\tau}^2 - T_{it}^2\right).$$

(iii) Identification of $(\beta_1, \mu_{\eta q}, \mu_{\eta \ell})$ comes from the three moments listed in equations (8) – (10).

(iv) Subtracting out variables that are known allows us to identify the joint distribution of the residual $\theta_i^w + \eta_{ij_i(t)} + \varepsilon_{it}$ across individuals. To identify the distribution of ε, we make use of Theorem 2.2 of Prakasa Rao (1992) which he attributes to Kotlarski (1967). We will refer to this as the Kotlarski theorem which We rewrite as:

[15] In practice this means that we observe each event at least once for each individual.

Suppose that X_1, X_2, and X_3 are independent real valued random variables. Define

$$\begin{aligned} Z_1 &= X_1 - X_3 \\ Z_2 &= X_2 - X_3 \end{aligned}$$

If the characteristic function of (Z_1, Z_2) does not vanish then the joint distribution of (Z_1, Z_2) determines the distributions of (X_1, X_2, X_3) up to a change of the location.

Consider using any two observations on the same job for a person (i.e. $j_i(\tau) = j_i(t)$), the residuals take the form $\left(\theta_i^w + \eta_{j_i(\tau)} + \varepsilon_{i\tau}, \theta_i^w + \eta_{j_i(t)} + \varepsilon_{it}\right)$. Applying the Kotlarski theorem we can identify the distribution of ε.

(v) Consider wages from the first period working and from the first period working following a layoff (which occurs at t^*). This gives us the joint distribution of

$$\left(\theta_i^w + \eta_{ij_i(0)} + \varepsilon_{i0}, \theta_i^w + \eta_{ij_i(t^*)} + \varepsilon_{it^*}\right).$$

Since the distribution of ε is already known, one can identify the joint distribution of $\left(\theta_i^w + \eta_{ij_i(0t)}, \theta_i^w + \eta_{ij_i(t^*)}\right)$ by forming the characteristic function of $\left(\theta_i^w + \eta_{ij_i(0t)} + \varepsilon_{i0}, \theta_i^w + \eta_{ij_i(t^*)} + \varepsilon_{it^*}\right)$ and dividing by the characteristic function of ε in the appropriate way. Apply the Kotlarski theorem again and we can identify of $F_\theta, F_{\eta 0}$, and $F_{\eta 1}$. One can use an analogous argument to identify $F_{\eta 3}$.

We have now shown identification of $\gamma, \beta, F_\theta, F_{\eta 0}, F_{\eta 1}, F_{\eta 3,}$ and F_ε. In showing identification of the remaining variables we will make use of the following argument.

Let t^* represent the date of the first employment spell following a nonvoluntary layoff. Let M_{it} be a generic random vector with $0 < t < t^*$. As a result M_{it} is independent of $\left(\varepsilon_{i0}, \eta_{ij_i(t^*)}, \varepsilon_{it^*}\right)$. Consider identification of the joint distribution of $\left(M_{it}, \theta_i^w, \eta_{ij_i(0t)}\right)$ from the joint distribution of

$$\left(M_{it}, \theta_i^w + \eta_{ij_i(0)} + \varepsilon_{i0}, \theta_i^w + \eta_{ij_i(t^*)} + \varepsilon_{it^*}\right).$$

Its corresponding characteristic function is:

$$\psi_{m,0,t^*}(\tau_1, \tau_2, \tau_3) = E\left(\left[\exp\left(i(\tau_1' M_{it} + \tau_2\left(\theta_i^w + \eta_{ij_i(0)} + \varepsilon_{i0}\right) + \tau_3\left(\theta_i^w + \eta_{ij_i(t^*)} + \varepsilon_{it^*}\right)\right)\right]\right)$$

where τ_1 is a vector (whose dimension is the same as M_{it}) and τ_2 and τ_3 are scalars. Note that ε_{i0}, ε_{it^*},and $\eta_{ij_i(t^*)}$ are independent of all other variables in the model so that

$$\begin{aligned}\frac{\psi_{m,0,t^*}(\tau_1,\tau_2,\tau_3)}{\psi_\varepsilon(\tau_2)\psi_\varepsilon(\tau_3)\psi_{\eta_1}(\tau_3)} &= E\left(\left[\exp\left(i(\tau_1' M_{it}+\tau_2\left(\theta_i^w+\eta_{ij_i(0t)}\right)+\tau_3\theta_i^w\right)\right]\right) \\ &= E\left(\left[\exp\left(i(\tau_1' M_{it}+\tau_2\eta_{ij_i(0t)}+\left(\tau_2+\tau_3\right)\theta_i^w\right)\right]\right)\end{aligned}$$

which is the characteristic function of $\left(M_{it},\theta_i^w,\eta_{ij_i(0t)}\right)$. Therefore the joint distribution of $\left(M_{it},\theta_i^w,\eta_{ij_i(0t)}\right)$ is identified.

(vi) Let $M_{it}=\left(X_{it},E_{it},T_{it},\ell_{it}\right)$. Since the joint distribution of $\left(M_{it},\theta_i^w,\eta_{ij_i(0t)}\right)$ is identified, the joint distribution of $\left(X_{it},E_{it},T_{it},\ell_{it},\theta_i^w,\eta_{ij_i(0t)}\right)$ is identified and therefore the distribution of ℓ_{it} conditional on $\left(X_{it},E_{it},T_{it},\theta_i^w,\eta_{ij_i(0t)}\right)=I_{it}^e$ is identified.

(vii) Let $M_{it}=\left(X_{it},E_{it},T_{it},1\left(\ell_{it}=2\right),1\left(\ell_{it}=2\right)w_{it+1}\right)$. From this the distribution w_{it+1} conditional on $\left(I_{it}^e,\ell_{it}=2\right)$ is identified. But since we are conditioning on X,E, and T, and since the distribution of ε is known, we can identify the distribution of η_{t+1} conditional on $\left(I_{it}^e,\ell_{it}=2\right)$. Thus $F_{\eta'}\left(\cdot;I_{it}^e\right)$ is identified.

(viii) Define M_{it} to be the state variables up until the point at which the person enters nonemployment. That is

$$M_{it}=\begin{cases}\left(X_{it},E_{it},T_{it},\ell_{it},e_{it+1}\right) & t<t_i^* \\ 0 & \text{otherwise}\end{cases}$$

Define $Y_{it}=\left(y_{i0},...,y_{iT}\right)$. Using the result above we know that we can identify this full joint distribution. Since the state space is limited, from this joint distribution we can identify the distribution of ℓ_{it} conditional on $\left(X_i,E_{it},T_{it},\ell_{it},\theta_i^w,\eta_{ij_i(0t)}\right)$ for $E_{it}\leq T$ and $T_{it}\leq E_{it}$.

(ix) Finally, consider identification of the job finding probability. This is somewhat more complicated because now $I_{it}^u=\left(E_{it},X_i,\theta_i^u\right)$. I need to put some more structure on the problem for θ_i^u to even be a meaningful object. Assume that this choice takes the form of a binary discrete choice model so that

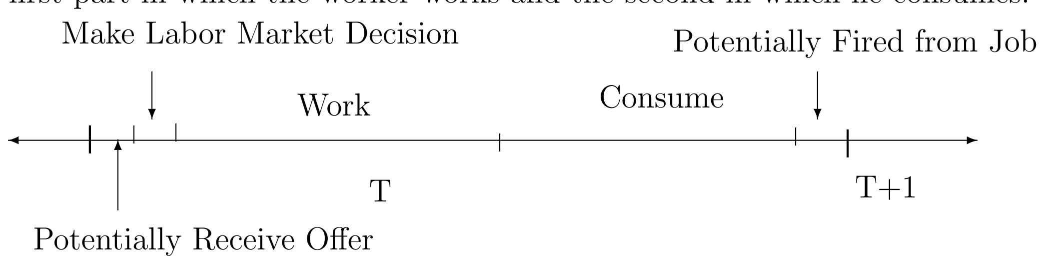
$$\Pr(s_{it+1}=0\mid\ell_{it}=1,I_{it}^n)=G\left(m(E_{it},X_i)+\theta_i^u\right).$$

With panel data on multiple spells this model would be identified if (E_{it}, X_i) were independent of θ_i^u using a sufficient normalization (see for example Cameron and Taber (1998)). In our model (E_{it}, X_i) is independent of θ_i^u conditional on θ_i^w. Thus one can use the type of argument above with M_{it} representing finding jobs after a nonemployment spell to condition on θ_i^w. Combining this with an argument about identification of the binary choice model (such as in Cameron and Taber, 1998) the job finding probabilities are identified.

A.2 An example of a Model of Floundering

In this appendix we present an example of a foundering model that justifies the approach we take. In a model using hyperbolic discounting, the "optimal" profile represents the turnover choices to which workers would pre-commit if there was a mechanism allowing them to do so. To the extent that workers differ from this optimal profile, their actual turnover choices can be considered floundering that occurs because they have time inconsistent preferences and no mechanism to constrain their future choices.

We assume that workers are risk neutral so that in a base model they behave as in the "optimal investment" case described above. Our goal is to model the notion that a worker might get mad at his boss and quit without thinking about the repercussions. Alternatively he might decide to stay in bed rather than coming into work on a particular day. To model this, we modify the timing of the model above so that each period is divided into two parts: the first part in which the worker works and the second in which he consumes.



Let ξ_{it} be a non-positive i.i.d. shock that a worker receives during the first sub-period if he stays on the same job. We allow for hyperbolic discounting as in Laibson (1997) to embody the idea of "without thinking about the repercussions." Specifically we allow for hyperbolic discounting between the first and second part of a period. When making labor market decisions at the beginning of period t the worker maximizes

$$\begin{aligned}
&E\left[\xi_{it} 1\left(j_{it}=j_{it-1}\neq 0\right)\right.\\
&\left.+\beta\left\{1\left(s_{it}=1\right)w_{it}+\sum_{\tau=t+1}^{T}\left(\frac{1}{1+r}\right)^{\tau-t}\left(\xi_{i\tau}1\left(j_{i\tau}=j_{i\tau-1}\neq 0\right)+1\left(s_{i\tau}=1\right)w_{i\tau}\right)\right\}\right].
\end{aligned}$$

If $\beta = 1$ this represents standard preferences. However, if $\beta < 1$ individuals have inconsistent preferences in the sense that the trade-off between ξ_{it} and w_{it} is different in the current period than it is when looking forward to future periods. A worker with a small β may react to a small change in ξ_{it} by quitting his job, even though that same worker would not want to quit this job when looking forward from $t-1$. In this case, if it were possible, the worker at time $t-1$ would choose to force himself to ignore the shocks ξ_{it} when making decisions. Interpreted in this way, the difference between actual behavior and "optimal" behavior shows how the worker's labor force outcomes would change if this type of pre-commitment were feasible.

Table 1

Summary Statistics

Male High School Graduates and Dropouts

First 10 Years Out of School

	Mean	St. Dev	Min.	Max.
Black	0.332	0.471	0	1
Hispanic	0.215	0.411	0	1
Education	11.239	1.102	9	12
Years in Data	9.716	1.344	1	10
Number of Jobs	6.205	3.77	1	26
Number of Times Fired	2.121	2.191	0	15
Number of Job to Job Transitions	1.933	1.981	0	14
Voluntary Switches to Nonemployment	1.331	1.416	0	8
Proportion of Time Working (monthly)	0.777	0.252	0.000	1.000
Sample Size: 2431 Individuals				

Table 2
Characteristics of Job Transitions
Male High School Graduates and High School Dropouts
First 10 Years Out of School

	Type of Job Transition		
	Involuntary	Job-to-Job	Job-to-Nonemp.
	N=4956	N=4968	N=3000
Log Wage Difference	-0.011	0.132	0.049
Black	0.348	0.272	0.378
Hispanic	0.238	0.206	0.220
Education (years)	11.029	11.225	10.924
Tenure (old job in months)	10.709	15.160	11.001
Experience (in months)	39.558	49.292	37.105
Length of Nonemployment Spell	5.869	-	9.939

Table 3: Panel 1

Estimates of Parameters of Model

(Standard Errors in Parentheses)

Estimates of Wage Equation	
Experience	0.047
	(0.010)
Tenure	-0.005
	(0.010)
Experience Squared/100	-0.070
	(0.030)
Tenure Squared/100	0.040
	(0.040)
Black	-0.110
	(0.021)
Hispanic	-0.017
	(0.023)
Education	0.055
	(0.008)
Estimates of Moments of Error Terms	
σ_ε	0.321
	(0.008)
σ_θ	0.253
	(0.010)
$\sigma_{\eta 0}$	0.252
	(0.016)
$\sigma_{\eta 1}$	0.222
	(0.008)
μ_θ	1.374
	(0.093)
μ_f	-0.017
	(0.029)
μ_ℓ	0.001
	(0.031)

Table 3 (cont.): Panel 2

Parameters of Job Transition Matrix

Multinomial Logit Coefficients

Variable	Fired From Job	Switch Jobs	Quit to Nonemployment
Black	0.203	-0.299	0.306
	(0.046)	(0.051)	(0.052)
Hispanic	0.109	-0.144	0.004
	(0.053)	(0.056)	(0.063)
Education	-0.153	-0.026	-0.232
	(0.017)	(0.020)	(0.019)
θ^w	-0.429	0.095	-0.412
	(0.160)	(0.133)	(0.166)
η	-0.315	-1.929	-0.884
	(0.209)	(0.187)	(0.229)
Experience/10	0.018	0.037	-0.101
	(0.020)	(0.026)	(0.024)
Experience Squared/100	-0.005	-0.004	0.004
	(0.002)	(0.002)	(0.002)
Tenure/10	-0.538	-0.146	-0.349
	(0.027)	(0.035)	(0.036)
Tenure Squared/100	0.027	0.000	0.012
	(0.003)	(0.004)	(0.005)
Intercept	-0.695	-3.393	-0.271
	(0.299)	(0.301)	(0.332)

Table 3 (cont.): Panel 3

Parameters for Finding New Jobs

Logit Coefficients

	Search From Fired	Find Job Immediately	Search From Quit
Black	-0.502	-0.382	-0.552
	(0.076)	(0.083)	(0.094)
Hispanic	-0.016	-0.197	-0.142
	(0.092)	(0.090)	(0.110)
Education	0.174	0.098	0.070
	(0.030)	(0.032)	(0.035)
Experience/10	0.056	0.078	0.147
	(0.032)	(0.036)	(0.039)
Experience Squared/100	-0.003	-0.003	-0.013
	(0.003)	(0.003)	(0.004)
θ^w	0.417	1.050	-0.595
	(0.265)	(0.206)	(0.274)
θ^{u*}	-0.780	-0.330	-1.093
	(0.045)	(0.069)	(0.050)
Intercept	-4.235	-3.275	-1.920
	(0.447)	(0.455)	(0.553)

Table 3 (cont): Panel 4
First part of Polynomial Terms
(ρ defined in equation 10)

	Degree 0	Degree 1	Degree 2	Degree 3
Intercept	-0.205	-0.938	1.994	3.765
	(0.100)	(0.321)	(0.248)	(0.467)
θ^w	0.072	0.455	-0.654	-1.753
	(0.088)	(0.21)	(0.170)	(0.245)
η	-2.182	-1.904	0.514	2.225
	(0.406)	(0.129)	(0.375)	(0.371)
$X'\delta$	1.000	1.000	17.255	39.715
	(-)	(-)	(17.281)	(40.591)
$\theta^w \times \eta$	0.976			
	(0.188)			
$\theta^w \times X'\delta$	6.160			
	(20.736)			
$\eta \times X'\delta$	18.572			
	(51.950)			

32