Wage Risk, On-the-job Search and the Value of Job Mobility

Kai Liu

Norwegian School of Economics *

This version: December 2011

Abstract

This paper shows that job mobility is a valuable channel in response to labor market shocks for employed workers. I construct a model of wage dynamics jointly with a structural dynamic model of job mobility. The key feature of the model is the specification of wage shocks at the worker-firm match level, for workers can respond to these shocks by changing jobs. The first result is that the variance of match-level shocks is large, and the consequent value of job mobility is substantial and decreasing in the cost of switching jobs. The second result is that true wage risk is more than twice as large as the wage variance observed after job mobility, which is what other papers in the literature have called wage risk.

JEL: D91, J31, J62

*Previously circulated under the title "Wage Risk, On-the-job Search and Partial Insurance". I am grateful to Robert Moffitt and Stephen Shore for guidance and support and to Peter Gottschalk for detailed comments. I also thank the participants at numerous seminars and conferences for helpful comments and discussions. All remaining errors are mine.
Email: kai.liu@nhh.no.

1 Introduction

Understanding how much idiosyncratic risk people face and how they respond to risk is an important topic for research. For most employed workers, wage risk is arguably the most important type of risk. There is an extensive literature analyzing individual's precautionary behavior such as savings and labor supply induced by idiosyncratic wage risk.[1] The implications of all these models depend critically on their having correctly identified wage risk, both its variation and persistence.

In most papers, wage risk is identified from the variance of wage residuals in a panel data model. In these papers, changes in properly defined wage residuals represent shocks. However, both the levels and the changes of wages can be endogenous and be the outcome of workers' choices. For instance, in a frictional labor market, people select better wages as they become available through on-the-job search. Furthermore, if there are shocks that are firm specific, a rational worker would respond to negative shocks by switching employers, thereby making both the degree of wage risk and the persistence of shocks different from those revealed in data. Moreover, with shocks mixed with endogenous choices, it is difficult to assess the true welfare cost of wage risk, to derive empirical implications of precautionary behavior, and to evaluate the consequences of government policy interventions. To understand true wage risk, it is essential to specify the sources of shocks and to model the individual's job choice together with the wage process.[2]

This paper makes two main contributions to the literature. First, it quantifies the value of job mobility as a channel of response to certain types of labor market risk facing employed workers. The value of job mobility depends critically on the cost of job change and the variation of wage shocks at worker-firm level. Second, the model is capable of recovering true wage risk that workers experience prior to job mobility. In the past literature, wage changes are usually observed after job mobility decisions and hence are different from the true shocks that occur prior to job mobility.

This is the first paper which studies the welfare value of job mobility as a mechanism for worker to respond to labor market shocks. The value of job mobility in this context builds upon two factors. First, I distinguish two sources of wage shocks in the wage equation: shocks at worker-firm match level and

[1] Among others see Deaton (1992); Carroll (1992); Gourinchas and Parker (2002) (precautionary savings) and Low (2005) (precautionary labor supply). See Heathcote, Storesletten, and Violante (2009) for a review.

[2] Recent papers by Low, Meghir, and Pistaferri (2010) and Altonji, Smith, and Vidangos (2009) also make important contributions in this direction. I discuss the differences between this paper and their papers in the next section.

shocks at the individual level which apply to all firms and matches. Contrary to shocks at individual level, shocks at match level do not mean permanent depreciation of individual's general productivity. Second, by modeling worker's job mobility decisions in response to labor market shocks, I show that match-specific wage fluctuations change the probability of job mobility. In a seminal paper, Topel and Ward (1992) find evidence that previous job-specific wage growth affects workers' job mobility decisions (holding the current wage and other observed characteristics fixed). However, they find this result "somewhat puzzling in light of our previous evidence that within-job wage growth approximates a random walk" (p.473). This suggests that one needs to estimate a stochastic wage process jointly with worker's job mobility choices, which is the direction taken in this paper.

I build and estimate a wage process jointly with a structural dynamic model of job mobility in an economy with search friction and job-switching cost. Switching costs are unobservable non-wage factors affecting worker's job mobility decision. I decompose the log wage into four independent and linearly additive components: a component which is predicted by personal characteristics, an individual component, a match component, and a transitory shock. The match component can be interpreted as job-specific human capital or idiosyncratic firm effect on wages.[3] In a labor market with search frictions, there is a distribution of firms offering the same worker different values of a match. Employed workers are motivated to search on-the-job and to choose a better match component of the wage as they locate other jobs over time. The match component and individual component follow parallel stochastic processes: each of them evolves from a permanent shock and a random growth factor. Shocks therefore represent permanent deviations from individual-specific wage growth profile. It is worth noting that the wages considered throughout this paper refer to real wages. Sticky nominal wages would show up as real wage cuts over time, and workers could be motivated to switch to other jobs that are willing to compensate for the cost of inflation. Similar to Topel and Ward (1992), I find strong empirical evidence that workers' mobility decisions are correlated with job-specific wage changes in the past.

The model is estimated by method of simulated moments using longitudinal data of young male workers from the 1996 panel of Survey of Income and Program Participation (SIPP). To separately identify the match component from the person component in the wage residuals, the model assumes that the match component is correlated with job mobility choices but the person component is not.

[3]Empirically it is infeasible to distinguish pure firm effect from pure worker-firm match effect without employer-employee matched data.

The key findings are the following: (1) Wage risk at the match (i.e., worker-firm) level accounts for the majority of the wage risk facing workers. (2) True wage risk, identified jointly from wage outcomes and mobility choices, is more than twice as large as the wage risk that is estimated using post-mobility wage information alone. (3) The welfare value of job mobility in response to match-level shocks is nearly 15% of lifetime expected utility. The value is decreasing in the worker's switching cost and increasing in the variation of match-level shocks. (4) The non-wage factor in the form of job switching cost is an important determinant of job mobility decisions. Switching costs are smaller for those who are married, college educated and possessing a house. (5) Unobserved individual heterogeneity (ability) explains a major portion of wage inequality at the beginning of work life. Over time, match-specific wages (built upon match-level shocks) becomes a dominating contributor to wage inequality. (6) The estimated mean return to tenure is negative and the mean return to experience is positive. There is strong evidence for heterogeneity in the return to tenure.

The rest of the paper proceeds as follows. Section 2 describes this paper's relation and contribution to the existing literature. Section 3 introduces the wage process and builds a parsimonious on-the-job search model. Section 4 introduces the data and presents empirical evidence from data which motivates this study. Section 5 discusses the estimation and identification strategy. Section 6 presents estimation results. Section 7 defines and quantifies true wage risk and the value of job mobility. It also discusses the role of job mobility on wage growth and inequality. Section 8 concludes, followed by a discussion of policy implications and future research.

2 Relation to the Existing Literature

There is a substantial literature assessing the magnitude and persistence of idiosyncratic wage risk using error component models. These studies typically find a random walk process for the permanent component and a mean-reverting transitory component.[4] Some papers also emphasize the importance of initial heterogeneity in wage growth (random growth model) (e.g., Haider (2001); Guvenen (2007)) The wage process in this paper builds upon these findings and incorporates all the features considered important in the literature.

[4] See, among others, Abowd and Card (1989); Baker and Solon (2003); Blundell, Pistaferri, and Preston (2008); Meghir and Pistaferri (2004); Moffitt and Gottschalk (2011); Jensen and Shore (2008).

A more recent line of work attempts to identify the sources of transitory and permanent income shocks. This paper relates to this developing literature, with a focus on estimating the effect from one particular source: job mobility. Two recent papers, Altonji, Smith, and Vidangos (2009) and Low, Meghir, and Pistaferri (2010), make important contribution to the literature by modeling earning dynamics and employment choices jointly. Low, Meghir, and Pistaferri (2010) estimate a wage process incorporating an individual's selection process between jobs and into and out of employment. Their estimates suggest that, once job mobility decisions are controlled for, the variance of permanent shocks is much lower. This suggests that what has been identified as permanent wage risk from typical error component model contains variability due to responses to shocks through job mobility. Altonji, Smith, and Vidangos (2009) construct a rich statistical model of earning dynamics from equations governing wage determination, hours of labor supply, job-to-job transition and transitions into and out of unemployment. They show that job mobility and unemployment, among other factors, play a key role in the variance of earnings over a career.

The current paper has several differences with these two studies, however. One important difference is that both Altonji, Smith, and Vidangos (2009) and Low, Meghir, and Pistaferri (2010) assume that the worker-firm match component of the wage does not vary over the duration of the job.[5] Within-job wage changes are assumed independent of worker's job mobility decision. Therefore, there is no match-specific wage risk except unemployment risk. One key feature of the current paper is to model wage dynamics within jobs and worker's selection across jobs. By doing so, it distinguishes wage risk that is particular to a job (worker-firm match) from wage risk applying to all jobs. In Section 4, I present evidence that job mobility is strongly correlated with past within-job wage changes. The estimation results indicate that modeling the evolution of match within jobs yields a very different picture of the wage risk facing workers. This paper also incorporates stochastic non-wage factors affecting job mobility, without which one risks biasing the estimated wage parameters.

A second difference is that Altonji, Smith, and Vidangos (2009) work with an econometric model without a utility maximization framework. While their descriptive statistical model may be attractive in many ways, their model does not permit one to estimate the welfare value of job mobility.[6] In terms

[5] Low, Meghir, and Pistaferri (2010) experiment with a random walk process of the match component with an individual fixed effect u_i as an alternative specification (in their "Robustness Checks" section), and conclude that it doesn't fit the data very well. However, they conjecture that some "combined aspects of stochastic specifications" may do better. In this sense, this paper advances their research agenda.

[6] For instance, they include other events such as health shocks and labor supply shocks in the earning process besides

of estimation strategy, Low, Meghir, and Pistaferri (2010) identify their wage process building upon a separately estimated job selection rule without using all the restrictions implied from the structural model. I adopt a more efficient estimation strategy by estimating the search model jointly with the wage process, since the search model implies structural selection process between jobs.

There is a growing literature trying to understand channels available for individuals in response to labor market risk.[7] A distinction is made between ex-ante and ex-post response to risk, where the former usually refers to insurance in anticipation of risk and the latter refers to reactions after shocks take place (Meghir and Pistaferri, 2011). This paper argues that, once the distinction between match- and person-specific shocks is made, job mobility is another useful channel that workers use ex-post in response to wage risk. This paper is also related to the partial equilibrium on-the-job search model in the literature.[8] A distinctive feature of my model is that there are two unobserved stochastic wage components evolving in parallel, yet under certain assumptions the decision rules can still be described by a set of reservation values. Moreover, this feature of the model provides an alternative within the on-the-job search framework to quantitatively match the extent of job-to-job transitions observed in the US labor market, where under plausible parameters, the basic on-the-job search model has failed to achieve.[9] It also provides a parsimonious way to explain wage cuts in job-to-job transitions. Job mobility with wage cuts has been difficult to reconcile in the standard search model because, with a stationary wage policy, the worker chooses to switch jobs only if there exists a job offering a higher wage.[10] In this paper, between-job wage cuts could arise out of three situations: a cut in the person-component of wage, a negative match-level shock and measurement error.

job mobility.

[7] Examples include Low (2005) (labor supply), Kaplan (2009) (within family), Blundell and Pistaferri (2003) (means-tested program), Gruber (1997) (unemployment insurance), Low and Pistaferri (2010) (disability insurance).

[8] Burdett (1978) is the pioneer work. See Mortensen (1986); Rogerson, Shimer, and Wright (2005) for a review.

[9] See Nagypal (2005) and Hornstein, Krusell, and Violante (2011) for further discussions.

[10] Postel-Vinay and Robin (2002) and Dey and Flinn (2005) rationalize this behavior through an on-the-job search with wage renegotiation between worker and current employer responding to outside offers. Hedonic models provide another explanation. Many structural estimations of search model (e.g. Wolpin (1992)) assume that observed wages contain measurement errors in order to produce positive likelihood of wage cut.

3 The Model

3.1 The Wage Process

The life-cycle wage process for an individual i employed by firm j in his labor market age t is:

$$\ln \widetilde{w}_{ijt} = \ln w_{ijt} + v_{it} \tag{1}$$

$$\ln w_{ijt} = Z_i'\beta + a_{ijt} + u_{it} \tag{2}$$

$$a_{ijt+1} = \begin{cases} a^l_{ijt+1}, & \text{if no job change between } t \text{ and } t+1 \\ a^o_{ij't+1}, & \text{if there is job change between } t \text{ and } t+1 \end{cases}$$

$$a^l_{ijt+1} = a_{ijt} + c_i + \eta_{ijt+1},\ a^o_{ij't+1} \sim N(0, \sigma^2_{a_0}) \tag{3}$$

$$u_{it+1} = u_{it} + \delta_i + \zeta_{it+1} \tag{4}$$

Assume that

$$E(\zeta_{it}) = 0, var(\zeta_{it}) = \sigma^2_\zeta \tag{5}$$

$$E(\delta_i) = \mu_\delta, var(\delta_i) = \sigma^2_\delta \tag{6}$$

$$E(v_{it}) = 0, var(v_{it}) = \sigma^2_v \tag{7}$$

$$\eta_{ijt} \sim N(0, \sigma^2_\eta), c_i \sim N(\mu_c, \sigma^2_c) \tag{8}$$

with orthogonality between all five of these error terms. $\ln \widetilde{w}_{ijt}$ is the observed real log hourly wage for worker i employed by firm j in labor age t and v_{it} is an error term combining a transitory component with measurement error (more on the latter below). Z_i is a $K \times 1$ vector of exogenous regressors of observed heterogeneity including a constant.[11] β is a $K \times 1$ parameter vector. For an employed worker, the log wage residual is decomposed into three components: an individual component u_{it}, a match component a_{ijt} between firm j and worker i, and the transitory shock v_{it}. The former two components evolve independently under two parallel stochastic processes, both with the same random walk, random growth, and transitory-component structure.

[11] The age profile of wages will be captured through the growth of random growth factors δ_i and c_i and through job selections as a result of on-the-job search. This is different from the literature treating wage growth as exogenous, where Z typically includes a constant and a quadratic in age and thus μ_δ is normalized to 0. See Section 5 for more discussions.

The individual component evolves over the lifecycle from an identically and independently distributed permanent random shock ζ_{it} and a random growth factor δ_i with mean μ_δ. The individual component measures the worker's general productivity regardless of his employer. It corresponds to the concept of permanent wage in the literature, which is usually thought of as representing return to skill or flow from human capital. The random growth factor δ_i has cross-sectional variance σ_δ^2. The heterogeneous growth in individual component of wage captures heterogeneous return to work experience, perhaps through differential learning ability to general skills or human capital investment.

Parallel to the individual component and prior to selection between jobs, the match component has a random walk, random growth, and transitory process. Let a^l_{ijt+1} be the latent match at $t+1$ prior to job mobility ("l" represents latent). It evolves from a random growth factor (drift) c_i with mean μ_c and cross-sectional variance σ_c^2 and a permanent shock η_{ijt}. I assume that the η_{ijt} are identically and independently distributed across firms, workers and time.[12] Unlike person-level shocks, I impose distributional assumptions on the match-level shocks and the random growth factor.

One interpretation of the match component is that it is an idiosyncratic firm effect which is a complement to individual productivity. From the perspective of human capital theory, the match component can also be regarded as job-specific human capital. The random growth factor c_i measures the individual-specific growth of match value for an employed worker, which can be thought of as return to job tenure or firm-specific human capital.[13] The shock to the match component then represents a worker-firm specific permanent deviation from the mean growth rate. This would happen, for example, when in a particular year the firm does not provide enough training to enhance worker's firm-specific skills (negative η_{it}), or it adopts a new technology that is complementary to worker's productivity (positive η_{it}). In general it consists of both a pure match-specific shock and a pure firm-specific shock, although without firm level data, distinguishing between these is not feasible. More broadly, the match component can be interpreted as any factor that affects the worker's productivity with the current firm but not after he leaves for other firms. The random growth factor and permanent shocks to the match

[12]Because the distribution of η_{ijt} is independent of firm, for simplicity, henceforth I drop the j subscript on η.

[13]This is an extension to the prototype model of wage determination in the literature which is used to estimate return to tenure (e.g. Abraham and Farber (1987); Altonji and Shakotko (1987); Topel (1991)), for here c_i is person-specific and correlated with job mobility. It turns out that there is substantial heterogeneity in c_i (Section 6). It would be interesting to model tenure effect as match-specific, so for the same worker, some jobs are expected to offer higher growth prospects but lower initial wage (compensating wage differentials). In a panel such as SIPP where the number of wage observations per person per job is small, it is difficult to identify wage level from wage growth from job-to-job transitions.

component are accumulating only over the current job tenure and will "vanish" after a job change.[14]

Flinn (1986) and Topel and Ward (1992) show the importance of the match component in explaining wage growth among young workers. Postel-Vinay and Robin (2002) demonstrate that, in an equilibrium on-the-job search model with firm and worker heterogeneity and wage renegotiation, log wages can be linearly decomposed into a worker-specific component and a firm-specific component closely interacting with labor market frictions. In that context, the match level shock specified here may result from a shock to worker's bargaining power or a renegotiation on wages following a credible outside offer to the worker. Due to data limitations, I take a more agnostic approach. The worker's wage is determined in a simple partial equilibrium on-the-job search model.

A job offer with match-specific wage a^o ("o" stands for "offer") is a random draw from a stationary offer distribution. I assume that it follows a normal distribution with mean zero and variance $\sigma^2_{a_0}$. Because of the growth profile in the person-component of wage (due to δ_i), the offered levels of wages would be mean-shifting with worker's labor market experience.[15] Offered matches are assumed uncorrelated with worker's individual wage component, and hence each firm has a constant return to labor technology and there is no sorting in the labor market.

When worker i receives an offer from firm j' at time t, prior to making a job mobility decision, the worker is perfectly informed of his general productivity u_{it}, match-specific productivity a^l_{ijt} if he chooses to stay and the value of the offer $a^o_{ij't}$.[16] At any time, workers have perfect information about their current match value, the expectation of future match values, and the distribution of the match component in the labor market, but information on other job locations and their associated match value must be obtained through search. The stochastic processes of match values guarantee that the search process does not eventually lead to a competitive outcome where all workers end up staying at the job of highest firm effect. I assume that none of the shocks to the u_{it} and a_{ijt} are anticipated by the worker so they represent wage uncertainty.[17]

Transitory shocks are identically and independently distributed across individual and time. The

[14]It is important to emphasize that the new accepted match would be positively correlated to the old match because of selection. It is only in this sense that firm-specific human capital may be partially transferable between jobs.

[15]A more difficult case is to allow offered match to depend on current and lagged match values, which is left for future research.

[16]Another set of search models develops the idea that the value of a match is not known when firm and worker meet but is updated ex-post as more information arrives. See Jovanovic (1979).

[17]This excludes the possibility that parts of these random shocks may be known to workers in advance. See Cunha, Heckman, and Navarro (2005).

transitory shock represents a wage shock with no persistence, at either the worker-firm match level or the person level. It also includes classical measurement errors on reported wages. I assume that transitory shocks affect wages *after* the mobility decision is made in each period, and therefore, given it is serially uncorrelated, it is unrelated to job mobility choices. This assumption simplifies the solution to the dynamic programming problem to be introduced in the following section. It is theoretically possible to allow transitory shocks to be serially correlated and to operate at both the person and match level.[18] In this case, both permanent and transitory match-level shocks enter into worker's information set prior to job mobility decisions. Besides increasing computational burdens, it is difficult to justify a homogeneous correlation restriction on the transitory component of wages applying to all workers' information sets. For the match-level wage process, the distinction between permanent and transitory is less important: permanent match shocks, albeit permanent from the view of workers, can be transitory ex-post if job-to-job transitions occur quickly. For these reasons, I have assumed that all shocks at the match level are permanent ex-ante in the worker's information set.

When a worker starts his work life, his initial wages are:

$$\ln w_{i0} = Z_i'\beta + a_{ij0} + u_{i0} \tag{9}$$

$$a_{ij0} \sim N(0, \sigma^2_{a_0}) \tag{10}$$

$$E(u_{i0}) = 0, var(u_{i0}) = \sigma^2_{u_0} \tag{11}$$

where u_{i0} is the initial individual wage component and a_{ij0} signifies the random match draw from the same job offer distribution at the beginning of work life.

3.2 The On-the-job Search Model

I now present a simple dynamic discrete choice model where individuals conduct on-the-job search given the wage process described previously. Workers begin employment in $t = 0$ by receiving a random job offer. Each job offer is characterized by a match value between the worker and the firm. At the beginning of each subsequent period, the worker makes the following discrete choice: move to a different job if an offer arrives, or stay with the current job possibly paying a different wage. On-the-job

[18] Many papers, e.g. Moffitt and Gottschalk (2011); Haider (2001); Meghir and Pistaferri (2004), show that there is some serial correlation over time in the transitory shocks. However, all these papers do not model worker's job-to-job selection. Job mobility is arguably the main contribution to transitory shocks (Gottschalk and Moffitt, 1994).

search is costless. However, when workers switch jobs, they have to pay a one-time switching cost. The switching cost captures unobserved non-wage factors that drive job mobility choice, which may include, for example, relocation cost.

I focus on the worker's job-to-job transition and ignore the worker's voluntary and involuntary transition to unemployment. Ignoring worker's involuntary transitions means there is no layoff risk. This assumption is made for two reasons. First, it is easy to show that having a single exogenous probability of match dissolution does not change worker's job-to-job selection process. Hence it does not affect the parameters identified from direct job-to-job transitions and wage outcomes. Second, from an empirical point of view, job-to-job transitions are the dominating phenomenon for workers in the US labor market compared with the transition from employment to unemployment. For working-age male workers, my calculation shows that job-to-job transitions happen more than twice as often as transition from employment to unemployment (voluntary and involuntary combined). Neglecting voluntary unemployment eliminates unemployment from worker's choice set. For the young male workers in our data (see below), their value of nonmarket time (e.g. value of leisure and home production) is arguably small. Combined with stochastic wage changes within job, these factors make unemployment a less attractive option. The variation of match-level shocks needs to be larger in order to induce worker to voluntarily quit job to unemployment. Nevertheless, in my directions for future research discussed at the end, I suggest that adding unemployment to the model would be a good extension.

Individual i maximizes the expected value of the discounted sum of a time-separable utility function

$$\max_{M_s, s=t,t+1,\ldots,T} \; E_t\left[\sum_{s=t}^{T} \Gamma^{s-t}(u(w_{ijs}) - M_{is}k_i)\right] \tag{12}$$

where Γ is the discount factor, T is the length of the decision horizon, and E_t is the expectations operator conditional on information available in period t. k_i denotes the one-time utility loss if the worker switches jobs ($M_{is} = 1$). The individual's utility function is assumed to be $u(w) = ln(w)$. As will become evident later, the functional form of the utility function is inconsequential to individual's job-to-job choice rules. Hours of labor supply is assumed exogenous and inelastic. The wage evolves according to the wage process specified before.

Figure 1 depicts the process of job mobility decisions. At the beginning of each period, the worker receives an offer from a different firm with probability λ^e. If he accepts the offer, his match component for this period will be the match value offered by the new firm. His (residual) wage in this period will be the sum of the offered match value, this period's individual component plus any unexpected transitory shock. If he rejects the offer, his wage paid by the current employer then adjusts to a new level to absorb the contemporaneous tenure effect, shocks to the individual and match components of wages, and the transitory shock. In the discrete time model, the worker has to sit out at least one period if he chooses to stay with his adjusted match value before sampling a wage offer from another firm.

Letting $S(j,t)$ represent the set of all state variables at time t, the value function for a worker i employed by firm j in period t is defined by[19]:

$$V_t^e(S(j,t)) = \begin{cases} V_t^e(S(j,t), M_t = 0), & \text{if no offer arrives in period t} \\ \max\{V_t^e(S(j,t), M_t = 0), V_t^e(S(j',t), M_t = 1)\}, & \text{if offer arrives from firm } j' \text{ in period t} \end{cases} \tag{13}$$

where

$$\begin{aligned} V_t^e(S(j,t), M_t = 0) =& \ln w_{jt} + \Gamma(1-\lambda^e)E_t\left[V_{t+1}^e(S(j,t+1))|S(j,t))\right] \\ &+ \Gamma\lambda^e E_t\left[\max\begin{Bmatrix} V_{t+1}^e(S(j,t+1), M_{t+1} = 0|S(j,t)), \\ V_{t+1}^e(S(j',t+1), M_{t+1} = 1|S(j,t)) \end{Bmatrix}\right] \end{aligned} \tag{14}$$

$$V_t^e(S(j,t), M_t = 1) = V_t^e(S(j,t), M_t = 0) - k \tag{15}$$

In principle, the set of state variables S includes all variables that provide information on the worker's current and future wages and mobility decisions. In our model, since the person match components evolve independently and linearly additive, the worker's problem can be re-normalized such that the state variables only include $\{a_{jt}, c, k\}$.[20]

For job mobility to take place, the expected value of a job offer must exceed the current value of the job plus the cost of switch. Let $h_t(a_{jt}; c, k)$ denote the reservation match in period t conditional

[19] I omit the person subscript i through the rest of this section.

[20] The intuition is as follows. Worker's mobility decision is based on $\Delta V = V_t^e(S(j,t), M_t = 0) - V_t^e(S(j',t), M_t = 1)$. Both the value function of moving and of staying contain $U(t)$ where $U(t)$ measures the expected life-time value of individual wage component and contributions from observable characteristics. $U(t)$ is not firm-specific and is independent of mobility choices (canceled out in ΔV) and hence state variables other than $\{a_{jt}, c, k\}$ do not affect worker's behavior.

on worker's type c and k. The following proposition characterizes worker's reservation match value for moving:

Proposition 1. *When job changes are costly, a worker's optimal strategy is to set a reservation match value $h_t(a_{jt}; c, k)$ where a worker chooses to move if and only if there is an offer such that $a^o_{j'} > h_t(a_{jt}; c, k)$. Furthermore, for all $t = 1 \ldots T-1$, $h_t(a_{jt}, k)$ satisfies the following properties: (1) $h_t(a_{jt}; c, k) > a_{jt}$ if $k > 0$ (2) $\frac{\partial h_t(a_{jt};c,k)}{\partial k} > 0$ (3) $0 < \frac{\partial h_t(a_{jt};c,k)}{\partial a_{jt}} < 1$.*

Proof. See Appendix A. $\square$

The first property says, in light of positive switching cost, the worker's reservation match value is always greater than the match value with the current firm (which is the reservation match when switching cost is zero). In a dynamic model where agents are forward-looking, the reservation match needs to include the expected discounted long-run compensation of the switching cost that has to be paid upon moving. The second property shows that the reservation match is monotonically increasing in the cost of switch. Switching cost influence the extent of labor market inefficiency besides search friction. The third property indicates that the reservation match is monotonically increasing in the quality of the current match and that the rate of increase is smaller as a_{jt} grows.[21]

An important insight from the last reservation match property is that, following a negative match shock, the worker's reservation match becomes lower than the reservation match without the shock. There is a set of wage offers that are acceptable after the match-level shock which would not have been acceptable without the match-level shock. This is how job mobility arises as a channel of ex-post response to wage risk. The value of job mobility depends on how the match-level shock affects the worker's job mobility decision, holding the reservation wage fixed at each period. In Section 7.1, I formally define and quantify job mobility as a means of responding to shocks in the labor market. Note that the welfare value of job mobility defined here is not the same as shutting down all job mobility, because with random job offers and on-the-job search, there is job mobility even if workers were not to

[21]The economic intuition for the second result is that a worker whose match is low expects more job changes in the future. For these workers, conditional on k_i and c_i, their optimal strategy is to set a larger gap between the reservation match and their match paid by the current firm. These workers are more likely to receive an acceptable offer in the future, and by setting the gap large, they avoid paying too much switching cost before reaching a high wage level. This is the same conclusion drawn from earlier papers analyzing the dynamic effect of switching cost in job mobility (Hey and McKenna, 1979; Van Der Berg, 1992). These papers assume that wages are constant within a given job and derive properties of the reservation wage from the steady state. This paper derives the implication of switching cost in a finite-horizon model with stochastic wage changes within jobs.

move because of match shocks. This discussion also highlights the economic importance of modeling the dynamics of the match-specific wage a_j and the person-specific wage u separately. Although both person- and match-level shocks represent wage risk, permanent shocks at the person level do not have any impact on worker's behavior. If match quality a_j is constant within jobs, then job mobility would not be a useful channel to act against wage shocks.

4 Empirical Evidence

4.1 Data and Summary Statistics

The data set I exploit is the 1996 panel of Survey of Income and Program Participation (SIPP). It is a four-year panel comprising 12 interviews (waves). Each wave collects comprehensive information on demographics, labor market activities and types and amounts of income for each member of the household over the four-month reference period. There are two main advantages of using the SIPP. One is that it has a short recall period, making it an ideal data set to study short-term employment dynamics that are very common among young workers.[22] The other advantage is that the SIPP contains a unique job ID for every job an employed worker had through the sample period. It records job specific wages and hours at each interview date (every four months), allowing researchers to obtain the precise wage changes at the time when job transitions take place. These features make it an attractive data set to study short-term job mobility and wage dynamics.

I focus on the primary job, which is defined as the job generating the most earnings in a wave. Although SIPP has monthly information on job changes and earnings, the time unit in the analysis of this paper is four months (a wave). This avoids the seam bias if we were using monthly variables. Real monthly earnings and the wage is derived by deflating the reported monthly earnings and wage by monthly US urban CPI. The reported hourly wage rate is used whenever the worker is paid by hour. For these workers, the real wage per wave is the mean of monthly real wage over the four months. For workers who are not hourly paid, their real wages are obtained by dividing real earnings per wave by reported hours of labor supply per wave.[23] Job change is identified from a change in job ID between

[22] In the selected sample, if a worker is observed to change jobs in a given calendar year, 19% of them would experience multiple job changes within the same calendar year. This means that job mobility observations at annual frequency understate the extent of job-to-job transitions by about a fifth.

[23] For each month, respondent reports hours of work per week and how many weeks worked. Monthly labor supply is

waves. If an individual is unemployed through the wave, no job ID would be assigned.

The original SIPP 1996 panel has 3,897,177 person-month observations. I drop females, full-time students, the self-employed, the disabled, and those who are recalled by previous employer after a separation. I trim the population whose real wage falls into the top and bottom 1% of the real wage distribution by wave. Because there is no unemployment state in the model, unemployed workers and people who have intervening periods of unemployment are dropped.

When a worker is interviewed in the first wave of SIPP, it is likely that he has already worked at a job for certain periods. In the first wave of SIPP, respondents are asked the starting date of the present job. I use this information to construct correct job tenure for workers with elapsed job duration when they are first sampled. Subsequently, the tenure of the present job in next wave is just the recoded job tenure plus one unless a job change is observed in the sample. In this way, the worker's job tenure and wage information is available throughout the sample period.

As explained in Section 5 below, I solve the initial condition problem by only using individuals who are observed from the first period of work life. As an approximation, I consider a job starting between age 20 and 26 as worker's first job in the life cycle.[24] This leads me to select workers whose calendar age is between 20-26 at the time when their present job started. I further restrict the sample to include young workers aged at or below 30 when they first enter the sample.[25] I then keep workers who are included in SIPP for at least eight waves, and construct a panel of 1,211 workers whose wages are observed for eight periods.[26]

Summary statistics are provided in Table 1. Table 2 reports the distribution of total number of observed job changes in the sample. The initial experience level refers to the labor market experience observed in the first observation period. Since the selected sample consists of young workers, nearly 50% of the workers switch jobs at least once in the four-year sample period. The extent of job-to-job transitions decreases monotonically with the workers' labor market experience. There is also some evidence that within-job wage growth declines with experience. To investigate the influence of observed heterogeneity on the pattern of job-to-job transitions conditional on experience, I estimate

calculated as hours per week$\times$(weeks worked/weeks in month)$\times$4.33

[24]It is notoriously difficult to determine the time when a sampled individual enters the labor market and starts employment. I believe this assumption is a reasonable approximation.

[25]So when a worker is sampled, the maximum possible elapsed job duration is $(30-19)\times 3 = 33$ periods.

[26]Note that the panel is essentially unbalanced and longer than eight periods since we are using job mobility histories of each worker.

a Tobit model of total number of job changes using all workers. Table 3 shows that, conditional on experience, there is strong evidence that workers who own a house make more job-to-job transitions.

4.2 Wage Growth and Job Mobility: Descriptive Evidence

The model features endogenous within-job wage change which is correlated with worker's job mobility decision. To evaluate this assumption, I first ask two sets of descriptive questions: First, what is the pattern of within-job wage growth and, in particular, how common are real wage cuts? Second, what is the empirical relation between within-job wage growth and subsequent job mobility, and between the level of wage and past mobility? Are workers who experience within-job wage cuts more likely to change jobs? Recognizing that the job-specific match is unobserved and that the job mobility decision is endogenous, the empirical evidence provided here does not carry any structural interpretation. Nevertheless, descriptive regressions in this section are useful benchmarks to evaluate the assumptions and implications of the model.

Figure 3A depicts the unconditional distribution of within- and between-job real wage growth. Wage growth is calculated as the change in log real wages every four months.[27] Two features of the picture are clear. First, between-job wage growth has larger variation than within-job wage growth, a feature consistent with the model because new job offers are random draws from a given distribution. Second, both within-job and between-job real wage cuts are very common. Around 45% of job-to-job transitions end up with wage cuts, and a little less than 50% of within-job real wage growths are negative. The majority of the wage cuts are small in magnitude. The median within-job real wage cut is merely 1.3% per period. There remains, however, a substantial portion of within-job wage growth showing significant drops. 15% of the workers report within-job real wage declines of 11% or more between waves. Wage cuts between jobs are much greater: the median between-job wage decline is about 20%. Measurement error may be an important contributor, as I discuss momentarily. Part of this could also be due to the stickiness of wages which are not immediately keeping up with a rising cost of living. Figure 3B shows the distribution of between and within nominal wage growth. While the majority of the workers experience nominal wage growth, the number of workers that had a nominal wage cut remains substantial.[28]

[27] Throughout this section, wage refers to real wage unless noted otherwise.

[28] Workers who are paid by hour experience less frequent nominal wage cut within jobs (yet remains sizable: about 20%

Columns (1a) and (2a) of Table 4 report descriptive regressions of log wage and mobility on lagged log wages and lagged job mobility. Current and two-period lagged mobility is not significant in column (1a), and the two-period lagged log wage is insignificant in column (2a), which could be due to measurement errors in observed wages (see below). The rest of the coefficients are significant and show the expected signs. These two regressions will be estimated on simulated data using the estimates from the on-the-job search model (see Section 6).

Next I investigate the empirical relation between within-job wage growth and worker's subsequent mobility choice. Specifically, suppose we have a worker employed by firm j at time $t-2$ and $t-1$. The primary question of interest is whether a worker whose within-job wage growth is low in period $t-1$ is more likely to move to another job in t. I estimate a probit model of job mobility on lagged wage growth. Table 5 reports the result. Column (1) shows the probit regression on one-period lagged within-job wage growth without any covariates. In (2) I add lagged job tenure and in (3) I further control for the two-period lagged level of the wage and lagged experience as explanatory variables. Columns (4) and (5) also control for two-period lagged within-job wage growth. In every specification, the coefficient on within-job wage growth in $t-1$ is negative and statistically significant. The coefficient on within-job wage growth in $t-2$ is also negative and significant in (5). This means that workers who experience smaller within-job wage growth are more likely to change jobs in the coming periods, even conditional on job tenure, labor market experience and the initial level of wage. This provides empirical support for us to model the dynamics of match and selection of match-level shocks through job mobility. The on-the-job search model also implies that job tenure is negatively correlated with probability of future job change, and workers whose levels of match are high are less likely to sample an acceptable offer and switch jobs. As expected, the estimated parameters on tenure and wage level are significantly negative. The experience parameter is insignificant conditioning on job tenure, which coincides with model's assumption that general human capital accumulation is independent of job mobility.

4.3 Measurement Error

In household surveys like SIPP, observed wages may be different from true wages because of reporting error. In our sample, measurement error in wages may come from two sources: from reported wages for

within job wage changes are wage cuts), than workers who are paid by salary.

those who are hourly paid, from reported earnings and/or hours for salary-paid workers. Assuming that these reporting errors are classical (i.i.d.), the extent of wage cuts documented previously may be an overstatement of true wage changes. Gottschalk (2005) uses earlier waves of SIPP and estimates that as many as three quarters of the observed decline in within-job nominal wages reflects measurement error. Our wage process incorporates classical measurement error in the transitory component of wages v_{it}. The estimated variance of transitory wage shock turns out to be fairly large (Section 6), suggesting that measurement error may have a large impact on the cross-sectional variation of observed wages.

Classical measurement error adds noise to true wages which would obscure the relation between job mobility and wage. If within-job wage changes are completely measurement error, there should be zero correlation between worker's job mobility choices and wage movement on the job. However, the probit regressions in Table 5 demonstrate a strong correlation between observed within-job wage changes and job mobility choices. With classical measurement error, the coefficient estimates on within-job wage growth term in Table 5 is biased downward, indicating that the true empirical relation between wage growth and job mobility is even stronger.

For nonclassical measurement error, studies typically find little effect of measurement error on earnings mobility and the covariance structure of earnings (Pischke, 1995). The more difficult question is whether measurement error is correlated with job mobility behavior. This could happen, for example, if the reporting error is smaller for workers with longer accumulated job tenure. Further empirical studies in this area are needed.

5 Identification and Estimation Strategy

5.1 The Structural Empirical Model

Motivated by the Tobit regression from Table 3, I begin by allowing the switching cost to vary with observable X:

$$k_i = \mathbf{X_i}'\boldsymbol{\Upsilon} \tag{16}$$

The vector X includes a constant, marital status, whether the individual owns a house, two education dummies(high school, at least some college) and a dummy on race. Υ is the vector of coefficients to be estimated.[29]

Recall that the wage equation for an individual i employed by firm j at labor market age t is:

$$\ln \widetilde{w}_{ijt} = \ln w_{ijt} + v_{it}$$
$$\ln w_{ijt} = Z_i'\beta + a_{ijt} + u_{it}$$

and the error terms evolve according to the following stochastic processes:

$$a_{ijt} = \begin{cases} a_{ijt-1} + c_i + \eta_{ijt} \equiv a^l_{ijt}, & \text{if } M_t = 0 \\ a^o_{ij't}, & \text{if } M_t = 1 \end{cases}$$
$$u_{it} = u_{it-1} + \delta_i + \zeta_{it}$$

where as previously defined, $a^o_{ij't}$ is the value of random offer, $a^l_{ij't}$ denotes latent match-level wage in period t prior to job mobility decision, and M_{it} is the observed mobility indicator, which equals to one when there is an acceptable offer. The job mobility decision can be formulated as a bivariate selection model:

$$J^*_{it} = (V^e_t(S(j',t), M_t = 0) - k_i) - V^e_t(S(j,t), M_t = 0) \tag{17}$$

$$J_{it} = \begin{cases} 1 \ if \ J^*_{it} > 0 \\ 0 \ elsewhere \end{cases} \tag{18}$$

$$O_{it} = \begin{cases} 1 \ with \ probability = \lambda^e \\ 0 \ with \ probability = 1 - \lambda^e \end{cases} \tag{19}$$

$$M_{it} = J_{it} \times O_{it} \tag{20}$$

where J^*_{it} is the offer-acceptance rule (choice) and O_{it} defines the outcome of a trivial offer-arrival process (chance) which is common across workers and independent of J_{it}. J^*_{it} is defined only over $O_{it} = 1$. Neither O_{it} nor J_{it} is observed by the econometrician; only the single indicator M_{it} is observed. The state variable vector contains $\{X_i, a_{ijt}, c_i\}$. Appendix B describes the solution method

[29] X_i is measured on the first observation date and is assumed time-invariant starting from the beginning of life.

to the value function in detail. The method uses Monte Carlo integration and an interpolation method to approximate the value function.

With zero switching cost, the offer acceptance rule can be simplified as:

$$J^*_{it} = a^o_{ij't} - a^l_{ijt} \tag{21}$$

The job selection rule is simply based on the difference between the offered match and the current match value, a result that can be easily extended from Burdett (1978). Evaluating the selection equation is equivalent to solving a static model for every period. I estimate both empirical models - the model without switching cost and the model imposing it. The preferred model is the model imposing job-switching cost.

5.2 Identification

The existing literature identifies wage risk using observed wages alone. I begin this section by illustrating why modeling job mobility decisions is necessary to identify the true wage risk. Suppose the log wage consists of only the match-specific component subject to permanent shocks. Figure 2 demonstrates two possible wage dynamics for a given worker. Prior to time t, the wage is a_0. At the beginning of period t, he suffers a permanent negative match specific shock η, and his new wage is $a_1 = a_0 - \eta$. The permanent wage drop considered here stems from a pure idiosyncratic firm effect and does not mean a depreciation of general individual productivity. In the absence of on-the-job search, his wage is expected to remain at a_1 for the rest of his working life.

I consider two scenarios. First, suppose a job offer valued a^o arrives at $t+1$ (left panel of Figure 2). Assuming a positive switching cost, if the new offer is greater than his reservation match $h(a_1)$, he would switch to the new job and earn a wage rate at $a_2 = a^o$. In this case, the wage increase from a_1 to a_2 results from an endogenous job mobility decision rather than wage risk, a point emphasized by Low, Meghir, and Pistaferri (2010). Moreover, by changing jobs, the worker manages to turn the initial permanent wage shock(η) into one that is effectively partly transitory and partly permanent. Only for a worker who remains at a_1 for a long time is the initial shock correctly identified. The ex-post(observed) persistence of the shocks depends on how quickly a worker could improve his match by changing jobs. Since the probability of job changes is inversely related to the quality of the contemporaneous match, the model implies that match-specific shocks would appear more persistent for workers of better match quality and less persistent for workers of lower match quality. When decomposing the variation of observed wage changes, the contribution from permanent shocks should then be larger for workers of higher match quality. This is in line with empirical evidence from the existing literature.[30] The right panel of Figure 2 depicts a second match dynamic in a similar setting. The only difference is that the worker is able to locate a better job within period t. If the worker takes the job, the observed wage rate in period t becomes a_2 which underestimates the magnitude of true wage shock. The observed average wage per period alone mitigates initial wage risk facing workers, as it is combined with worker's response to latent shocks. The variance of permanent match-level shocks, σ_η^2, measures wage risk prior to job mobility.

Our preferred model is parsimoniously characterized in terms of the parameter vector[31]

$$\Theta = (\beta, \lambda^e, \sigma_\delta^2, \sigma_c^2, \sigma_{u_0}^2, \sigma_{a_0}^2, \sigma_\eta^2, \sigma_\zeta^2, \sigma_v^2, \mu_c, \mu_\delta, \Upsilon) \tag{22}$$

Coefficients (β) on observable vector Z_i are identified by the initial wage assumption. Since the wage residual observed in the first period of work life is assumed exogenous (no selection takes place prior to the beginning of work life), an OLS regression of observed wages in the first period work life produces consistent estimates of β.

The preceding discussion demonstrates that job mobility choice invokes two selection equations: one is the trivial and exogenous offer-arrival process which does not depend on any covariates, and the other is the job-selection rule. In the full model with switching cost, the job-selection rule can be thought of as a reduced-form equation on X_i, which is analogous to the choice equation in the classic Heckman selection models. Therefore, the switching cost parameters (Υ) are identified by exclusion restrictions. The excluded state variables, which affect the decision to select between jobs but do not enter the wage equation, include marital status and house ownership. As a sensitivity test, I also estimate the model assuming zero switching cost, whose identification does not depend on the exclusion restriction. The

[30] For example, taking estimates from Table I and III of Meghir and Pistaferri (2004), a simple calculation shows that for college graduates, the variance of permanent shock account for 67% of variance of (unexplained) earnings growth, but the number drops to 27% for high school graduates and 20% for high school dropouts. This is consistent with implications from the model, if one believes that more educated workers acquire job-specific skills quicker and build up a higher match on average.

[31] The discount factor Γ is held fixed at 0.97.

estimated wage parameters are similar (see Section 6).

I partition the rest of the parameters into two sub-vectors: $\Theta^m = \{\lambda^e, \sigma^2_{a_0}, \sigma^2_\eta, \sigma^2_c, \mu_c, \mu_\delta\}$ and $\Theta^u = \{\sigma^2_{u_0}, \sigma^2_\zeta, \sigma^2_v\}$. Θ^m contains parameters governing the evolution of the match-level wage and the search process. Θ^u includes parameters determining the person-level wage process. The distinction between these two sets of parameters is that the former determines job mobility histories, while the latter is independent of job mobility. Appendix D shows that the covariance of wages and mobility, autocovariance of mobility, mean of mobility and the wage only depend on Θ^m. Θ^m is constant over time. With observations from sufficient periods, these moments overidentify all the parameters in Θ^m.[32] A widely recognized difficulty in estimating search models is that rejected offers (wage offers below the reservation wage) are unobserved. Hence the normality assumption on the offer distribution is necessary to recover the mass below the reservation wage (Flinn and Heckman, 1982). Analogously, extremely bad shocks to a_{ijt} and extremely low c_i are not observed if workers are able to switch jobs very quickly. Distributional assumptions on the shocks and random growth factor in the match-level process are essential in evaluating the mass at the bottom of the distribution.

Given Θ^m, Θ^u can be identified using moments based on the variances and autocovariances of wages alone, like most papers in the wage dynamics literature. No distributional assumptions are required to identify parameters governing the person-level wages. σ^2_δ and σ^2_ζ imply that the autocovariance of person-level wages grows nonlinearly and linearly, respectively, with time locations (see formula in Appendix D). Hence they are identified by fitting a quadratic and linear trend on autocovariance (Guvenen, 2007). Transitory shocks, which are i.i.d by assumption, are identified from the variances. Initial individual heterogeneity $\sigma^2_{u_0}$ is identified from the variance of wages at the beginning of life.

Lastly, I provide some intuition for identification, for the parameters in Θ^m. As is evident from equation (11) in Appendix D, the mean of return to experience (μ_δ) is identified by fitting a linear trend in the observed wage residuals over life. Within-job wage growth over life identifies the return to tenure and the combination of match-level uncertainty (σ^2_η) and heterogeneity in return to tenure (σ^2_c).[33] In a model without match-level shocks and heterogeneous growth factor c_i, within-job wage growth is exogenous and constant over life. In the data, however, the empirical evidence suggests

[32] Note that, in estimation, the parameters are estimated jointly imposing all the moments simultaneously.

[33] Within-job wage growth alone and first moment of wage in general do not separately identify σ^2_η and σ^2_c because they are linearly additive in $\sigma^2_{a^l_t}$. Intuitively, at the mean level, random growth factor c_i can be thought of as a special match-level shock drawn from a given distribution, except that the draw is accidentally the same for all periods.

that within-job wage growth slows down with experience (Table 2). By modeling the dynamics of the worker-firm match and job mobility, the model is able to capture the decline in within-job wage growth over time. Figure 4 demonstrates that, holding everything else constant, different σ_η^2 (or σ_c^2) changes the slope and curvature of within-job wage growth profile over time and different μ_c only changes its intercept. Therefore, the slope and curvature of the fitted profile of within-job wage growth identifies the combination of σ_η^2 and σ_c^2, and, given μ_δ, its intercept identifies μ_c.

To separately identify σ_c^2 and σ_η^2, one needs to rely on the autocovariance function of job mobility. To see this, Figure 5 plots the simulated autocovariance of mobility at different lags, given three combinations of σ_η^2 and σ_c^2 while holding the rest of the parameters fixed. While both greater σ_η^2 and σ_c^2 lead to an increase in the covariance of mobility across different lags, the relative effect from σ_η^2 quickly deteriorates with the length of the lag. Hence σ_c^2 can be separately identified from σ_η^2 using the autocovariance of mobility at sufficiently long lags.

Two remaining parameters, offered match heterogeneity ($\sigma_{a_0}^2$) and offer arrival probability (λ^e), are identified from the covariance of lagged wage and job mobility and the mean of job mobility. Specifically, $\sigma_{a_0}^2$ is identified by $E(r_0|M_1 = 1)$, which is a function of $\sigma_{a_0}^2$ but not λ^e (see equation (13) of Appendix D). Given $\sigma_{a_0}^2$ and the normality assumption on the distribution of offered matches, λ^e can be identified from the probability of switching jobs.

5.3 The Initial Condition Problem

Since SIPP is a short panel, it is typical that some workers have left-censored job histories when they are observed in the first wave of SIPP. For these workers, their first observed wages are endogenous which leads to an initial condition problem (Heckman, 1981). Recall from Section 4.1 that the SIPP contains information on the starting date of worker's present job when he is first sampled, and this information is used to select a sample of workers whose complete job histories are known from the beginning of work life. Let us define p ($p = 1, ..., P$) as the observation period for a worker in SIPP, and let τ represent the elapsed job duration in the first observation period ($p = 1$). Since we know that in our sample a worker's present job is his first job, p and τ together map into a unique work life period t: $p + \tau = t$.

Assume that wages in the first period of work life are exogenous. Then the initial condition problem can be solved by simulating the model starting from the beginning of life cycle and examining the desired statistics conditional on worker's past mobility choice. Specifically, we observe a worker's mobility choices starting from the beginning of work life: $M_0 = M_1 = ... = M_\tau = 0$. Starting from work age $\tau + 1$ through $\tau + P$, we observe both his mobility choice and wage information. Recognizing that the match value in the first period of life is an exogenous random draw, we can simulate job mobility and wage histories from the beginning of work life until $\tau + P$. Simulated statistics (e.g. mean wages) in period $\tau + p$ conditional on $M_0 = M_1 = ... = M_\tau = 0$ are consistent estimators for the observed statistics in observation period p.[34]

5.4 MSM Estimation

It is difficult to accurately estimate all elements of Θ in one step for computational reasons. I employ a two-stage estimation procedure. In the first step, I estimate the earnings regression and produce a consistent estimator of β and the log wage residual for each worker in each observation period in the sample. I select workers whose initial wage on their first job is observed[35]. As discussed previously, regressing it on observed personal characteristic Z yields a consistent estimator of β - the coefficient vector on Z. The vector Z includes a constant, education dummies, race, a polynomial in calendar age, and interactions between age and education dummies.[36] Consistent with the majority of findings from Mincerian earning regressions, the R^2 in the first stage regression is low ($= 0.17$), leaving the bulk of initial wage variation to unobserved heterogeneity in the match- and person-level wage components. For the remaining analysis I work with predicted log wage residuals denoted by r_{ip}.

The remaining parameter vector, denoted by θ, is estimated by method of simulated moments (MSM). For each worker in the sample, we observe a vector of mobility choices M_p and a vector of wage residuals r_p. From these two vectors, we can derive the empirical moments $s_i(M_{p_1}, r_{p_1}, M_{p_2}, r_{p_2})$

[34] If the worker's present job (at the time he is sampled) is not his first job, this approach is also valid provided we know exactly when the job change occurs. SIPP does not contain this information. In other words, the key is to have worker's complete pre-sample history of job mobility.

[35] 672 workers satisfy this criteria.

[36] *Age* here captures the effect of potential labor market experience on worker's initial wage. When calculating predicted residuals for each worker, *Age* is held fixed at the calendar age when he enters the labor market. By doing so, the age profile of wage residuals will be accounted for through mechanism of the model: through mean return to tenure (μ_c), return to experience (μ_δ) and worker's selection into better jobs.

for any two observation periods p_1 and p_2 such that $1 \leq p_1 \leq p_2 \leq P$:

$$\begin{aligned} s_i(M_{p_1}, r_{p_1}, M_{p_2}, r_{p_2}) =& [M_{ip_2}, r_{ip_2}, (r_{ip_2} - \overline{r_{p_2}})(M_{ip_1} - \overline{M_{p_1}}), (M_{ip_2} - \overline{M_{p_2}})(r_{ip_1} - \overline{r_{p_1}}) \\ & (r_{ip_2} - \overline{r_{p_2}})(r_{ip_1} - \overline{r_{p_1}}), (M_{ip_2} - \overline{M_{p_2}})(M_{ip_1} - \overline{M_{p_1}})]' \end{aligned} \tag{23}$$

where $\overline{r_p}$ is the population average of r_{ip} in period p; M_{ip} denotes the binary mobility choice and $\overline{M_p}$ is the population average of M_{ip} in period p.

The empirical model to be estimated is:

$$s_i(M_{p_1}, r_{p_1}, M_{p_2}, r_{p_2}) = f(\theta; X_i, \tau_i, p_1, p_2) + \varepsilon_i \tag{24}$$

Function f is the expected value of $s_i(M_{p_1}, r_{p_1}, M_{p_2}, r_{p_2})$ as implied from the model. It is important to note that the predicted moment f depends on τ_i (elapsed duration of the first job at $p = 1$) because of the left-censoring problem in the sample. τ_i, p_1 and p_2 map into two unique life periods t_1 and t_2. Conditional on observed state variables, the theoretical moments are functions of life period t_1 and t_2 conditioning on $M_0 = M_1 = ... = M_{\tau_i} = 0$.

The function f is complicated to compute analytically because, in the presence of endogenous selection on the match process, the distribution of the state variables at any given life period, $F(s(j,t); \theta, X_i, Z_i, \tau_i)$, is difficult to evaluate. I choose to approximate it by its simulated counterpart:

$$\hat{f}(\theta; X_i, \tau_i, p_1, p_2) = \frac{1}{S} \sum_{s=1}^{S} f(\theta; \hat{\nu}_s, X_i, \tau_i, p_1, p_2) \rightarrow f(\theta; X_i, \tau_i, p_1, p_2) \tag{25}$$

where, after taking the mean of the simulated sample, $\hat{f}(\theta; X_i, \tau_i, p_1, p_2)$ becomes the simulated moment vector containing the mean, variance and covariance of earnings and mobility at two work-life periods t_1 and t_2 which $\{\tau_i, p_1, p_2\}$ map into. $\hat{f}(\theta; X_i, \tau_i, p_1, p_2)$ converges to $f(\theta; X_i, \tau_i, p_1, p_2)$ as the number of simulations S becomes large. $\{\hat{\nu}_s\}_{s=1}^{s=S}$ is a sequence of random variables that are identically and independently distributed. It consists of sequence of job offer draws a^o, shocks to match component η and shocks to switching cost from all periods, and a vector of person-specific tenure effect c_i. With $\hat{\nu}$, the model is able to simulate S job histories for each individual.[37] Person-component variables are

[37]Vector $\hat{\nu}$ is held fixed across individuals of the same type because its distribution does not depend on i. A usual caveat to this frequency-type simulator is that the mobility function is non-smooth in the parameters, introducing difficulties for

additive in the match-component and do not affect the mobility rules. Therefore, person-level shocks are not simulated and they enter in $\hat{f}$ with an analytical expression. For example, one of the theoretical moments is $var(r_{it})$ consisting of a linear combination of $var(u_{it})$ and $var(a_{ijt})$. As the person-level wage u_{it} is independent of the job mobility decision, the former can be expressed as function of model parameters: $var(u_{it}) = \sigma^2_{u0} + t_2^2\sigma^2_\delta + t_2\sigma^2_\eta$. The latter, however, can only be evaluated using Monte Carlo simulations. Details of the simulation procedure are described in Appendix C.

Assume we have a balanced panel at hand for notational convenience. Taking sample averages, the LHS of equation (24) becomes the mean, variance and covariance of mobility and wage residuals between any two observation periods in the sample:

$$[E(M_{ip_2}), E(r_{ip_2}), cov(r_{ip_1}, r_{ip_2}), cov(M_{ip_1}, M_{ip_2}), cov(r_{ip_1}, M_{ip_2}), cov(M_{ip_1}, r_{ip_2})]'$$

The vector of simulated moments for any given two observation periods is:

$$\mathbf{g}(\theta; p_1, p_2) = \mathbf{s}(M_{p_1}, r_{p_1}, M_{p_2}, r_{p_2}) - \frac{1}{N}\sum_{i=1}^{N} \hat{f}(\theta; X_i, \tau_i, p_1, p_2) \tag{26}$$

where N is the number of workers in the panel, $\mathbf{s}$ is the sample average of $s_i(M_{p_1}, r_{p_1}, M_{p_2}, r_{p_2})$. Since we calculate f for every worker that we observe at p_1 and p_2, effectively, the simulated population has the same distribution of X, Z, τ as in the observed sample.

Let $\mathbf{g}(\theta)$ be a vector consisting $\mathbf{g}(\theta; p_1, p_2)$ at all possible combinations of p_1 and p_2. The size of vector $\mathbf{g}(\theta)$ is $M \times 1$. The total number of moments is $M = (P+1)(2P-1)$. I choose $P = 8$, leading to a total of $M = 135$ moments used.[38] The goal of the second stage estimation is to find θ which minimizes:

$$\mathbf{g}(\theta)'\mathbf{W}\mathbf{g}(\theta) \tag{27}$$

where $\mathbf{W}$ is an M by M weighting matrix. The weighting matrix is chosen as a diagonal matrix such

gradient-based optimization method. I smooth M_{it} by $\Phi(\frac{J^*_{it}}{h})$, where Φ is the standard normal cumulative distribution function and h is a smoothing parameter. When $h \to 0$, the approximation converges to the frequency simulator.

[38]Beyond eight sample periods, the sample size is insufficient to yield reliable estimates of empirical moments.

that $\mathbf{W} = A^{-1}$. A is a diagonal matrix whose elements on the main diagonal are given by

$$A_{kk} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{s}_i - \mathbf{s})(\mathbf{s}_i - \mathbf{s})'$$

Therefore, the main diagonal of $\mathbf{W}$ contains the inverse of the variance of corresponding elements in $\mathbf{s}$. Given the set of moments, this diagonally-weighted minimum distance estimator is more efficient than the equally-weighted estimator where W is an identity matrix. Intuitively, this weighting matrix allows us to adjust the moment conditions such that they all enter the objective function as being roughly similar in scale.[39] It also avoids the large small-sample bias if the optimal weighting matrix were used, which is primarily related to the off-diagonal elements in the optimal weight matrix (see Altonji and Segal (1996)).

6 Estimation Results

The estimated parameters are presented in Table 6. The first column contains estimates for our preferred model with switching cost. Wage risk at the worker-firm match level is the dominating risk facing employed workers: the variance of match-level shock (σ_η^2) is $\frac{0.986}{100}$, which is more than two orders larger than the variance of the person-level wage shock ($\sigma_\zeta^2 = \frac{0.005}{100}$). It implies that on-the-job search could potentially respond to a large fraction of permanent wage risk. The variance of the transitory shock is large, which is consistent with previous discussion that measurement errors could be responsible for substantial wage variations.

Turning to the random growth factors, I find that the point estimate of return to tenure (μ_c) is about -0.8% per period (four months) and the estimated return to experience (μ_δ) is 1.2%. There is strong support for heterogeneous return to tenure and return to experience. The model presented in this paper contains heterogeneous return to tenure and permanent match-specific shocks, both of which are absent in reduced-form models of the return to tenure.[40] Permanent match-level wage shocks and

[39] The specified weighting matrix is essential here, because the covariance between wage and mobility is much smaller than the variability of wage. I also estimated the model using an identity weighting matrix, which is equivalent to minimizing the sum of squared residuals $\sum_{i=1}^{N}\varepsilon_i^2$. The estimated match heterogeneity and offer arrival rates are severely downward biased, although the rest of the parameters are similar. The full results are available upon request.

[40] See Altonji and Williams (2005) for a reassessment of this literature. In reduced-form estimations of return to tenure, any shock to match component is assumed transitory and therefore does not relate to turnover behavior. Also, if return to tenure is heterogeneous, on-the-job search implies that workers whose returns are low tend to switch jobs at a faster rate, generating a positive relation to observed tenure. This is likely to produce a positive source of bias to existing estimators.

job mobility alone can generate sufficient positive wage growth over time, because workers are able to preserve good match shocks and move away from bad match shocks by job mobility. The negative μ_c coincides with Nagypal (2005), who shows that one needs to have a decreasing value of match quality over the job tenure in order to match the high rate of job-to-job transitions in the data. Our story is similar here: if there is no depreciation in job-specific skills (i.e. μ_c is positive), the model implies a quick decline in the probability of moving over time. This would be inconsistent to the high rate of job-to-job transitions we observe even for workers.

The estimated heterogeneity of the individual wage component at the start of life is much larger than the initial match heterogeneity, suggesting that the individual wage component is essential to match both the extent of job-to-job transitions and the associated dispersion in wages.[41] Examining the switching cost parameters, I find that the estimated switching cost is moderate for the sample under study. The magnitude of estimated $\gamma^{constant}$ is more than one-third of the mean log wage in the first period of life. For this sample of young male workers, the switching cost is smaller for workers who have a college degree, own a house, and for married and white workers.[42] The subsidy from spouses to job search could reduce the mobility costs of married workers. Young workers who own a house may be more likely to change jobs when there is a wage fall, perhaps because they have to make mortgage payments.

What happens if shocks to the worker-firm match and job switching cost are ignored? This corresponds to the assumption made in Low, Meghir, and Pistaferri (2010) and Altonji, Smith, and Vidangos (2009), where the worker's mobility choice is solely based on the value of initial match. Column (3) presents the estimated parameters by assuming constant matches within jobs. We see a large increase in the estimated variance of permanent shock (from $\frac{0.005}{100}$ to $\frac{0.176}{100}$). A large proportion of wage fluctuations that is in fact specific to a worker-firm match has been identified as permanent shocks that will persist across all jobs. In Section 7, I discuss the implications of this finding for the true wage risk facing workers. In column (4), I estimate a canonical wage process by neglecting the match-specific wage altogether and hence disallowing the worker's selection between jobs.[43] This has been a standard

[41] Bils, Chang, Kim, and Hall (2009) and Hornstein, Krusell, and Violante (2011), show that match heterogeneity alone is insufficient to produce both realistic wage dispersion and unemployment fluctuations at the same time.

[42] Note that the offer arrival process is homogeneous across workers. The switching cost parameters will be contaminated if the offer arrival rate also varies across worker's education, marital status, race and house ownership (e.g. because worker's search effort varies). One needs to explore other exclusion restrictions to separately identify the offer arrival process (J) and the job selection process (O).

[43] Since wage is assumed exogenous (no selection between jobs), the first stage regression is run using observed wages

in estimating wage uncertainty in the labor and macroeconomics literature. Compared with estimates from the model taking job mobility into account (but assuming constant match) in column (3), the permanent wage uncertainty nearly doubles and the variance of transitory shock increases by a third, from 0.036 to 0.046. These results are consistent with findings in Low, Meghir, and Pistaferri (2010). If we ignore the worker's selection between jobs, about 50% of the identified permanent wage uncertainty stems from worker's endogenous job mobility choice.

The identification of the full model hinges on exclusion restrictions. As a sensitivity test, I also estimate the model without switching cost, where job mobility is only influenced by wage differences. Column (2) presents the results. Comparing the first and second column, we find that the estimated parameters are qualitatively very similar. The variance of match shock is smaller, yet it remains very large relative to the variance of person-level shock. When non-wage factors are disallowed to influence job mobility, we also obtain a decline in the heterogeneity in offered match values. Unobserved non-wage factors raises the worker's reservation wage in general, which would presumably drive up the variation of match-level shock and offered match values in order to match the same extent of job-to-job transitions and wage growth from data.

To evaluate the fit of the model, I simulate 20 careers for each worker in the sample using the parameter estimates for the switching-cost model. Therefore I simulate a total of 24,220 careers and the simulated population has the same distribution of observed characteristics X and Z. I then truncate the careers according to the empirical distribution of left-censored job spells τ. The final simulated sample contains 8 observation periods, whose joint distribution of X,Z and τ matches the SIPP sample. Next, I run the wage and mobility regression on the simulated sample and compare the regression relationship between wage and mobility with those of the estimates from SIPP (Table 4). The estimated coefficients (column (1b) and (2b)) are very close to the estimates from SIPP. The only difference is the coefficient on one-period lagged mobility in the wage regression, where the model implies a negative correlation between lagged mobility and current wage. Overall, I conclude that the model is able to match closely the dynamic relationship between job mobility and wage in the SIPP sample.

over all periods on personal characteristics including labor market experience. Transitory shocks are assumed i.i.d. The weighting matrix used is an identity matrix.

7 Implications of the Model

7.1 Value of Job Mobility in Response to Wage Shocks

Match-level wage shocks affect the probability of job mobility for a given worker-firm match. To quantify the value of job mobility as a way of responding to the match-level wage shocks, I calculate the life-time expected utility of a worker holding the reservation wage fixed in every period at the level before match-level shock takeing place in that period (thereby disallowing job mobility to respond to match-level shocks). Mathematically, let $V_0^{e'}(S(j,0))$ be the value at the beginning of life in the hypothetical environment. $V_0^{e'}(S(j,0))$ is evaluated with respect to a modified conditional density function $f'(a_{ijt}|a_{ijt-1},\eta_{it})$ in every period:

$$f'(a_{ijt}|a_{ijt-1},\eta_{it}) = g'(a_{ijt}|M_{it},a_{ijt-1},\eta_{it})h'(M_{it}|a_{ijt-1},\eta_{it}=0) \tag{28}$$

where η_{it} is absent in the h' density but not in the g' density. Conditioning on the worker's type and history of match draws up to $t-1$, h' measures whether the worker would have moved in the absence of match-level shocks. Then, the value of job mobility at the beginning of life is given by

$$\Delta V_0 = V_0^e(S(j,0)) - V_0^{e'}(S(j,0)) \tag{29}$$

where $V_0^e(S(j,0))$ is the life value with free job mobility responding to match-level shocks. ΔV_0 measures the welfare value of job mobility in response to match-level risk. The higher the switching cost is, the closer is the $h'(M_{it}|a_{ijt-1},\eta_{it}=0)$ density from the $h(M_{it}|a_{ijt-1},\eta_{it})$ density and the smaller is the ΔV_0.

Table 7 shows the calculated value of job mobility. Panel A shows that the welfare value of job mobility (ΔV_0), when represented as a percentage of the value a job V_0^e, is close to 15% for all types of workers. The value of job mobility is monotonically decreasing in the cost of job changes: workers whose switching cost is low (e.g. married, do not own a house, college graduate, white workers) enjoy larger values of job mobility than those whose cost of switching employer is high (e.g. single, do not own a house, high school graduate, non-white workers). When switching cost is large, job mobility is a less useful tool to act against negative match-level wage shocks. Clearly, when switching cost is infinite, ΔV_0 should be equal to zero. Panel B suggests that, holding the switching cost fixed, the value of job mobility is increasing in the variances of match-level shocks. When wage variations in the match-component is high (like the one estimated earlier in this paper), job mobility is an extremely valuable channel to act against negative wage shocks.

7.2 Wage Risk Prior to Job Mobility

True wage risk corresponds to the wage risk prior to making the job mobility decision in each period. In our model, the true wage risk is the sum of the variance of the person- and match-level shocks ($\sigma_\zeta^2 + \sigma_\eta^2 = \frac{0.829+0.005}{100}$), which is more than twice as much as the variance of permanent shock identified from a canonical wage process where the firm-specific wage is neglected ($\frac{0.321}{100}$). When the switching cost is ignored, the estimated true permanent uncertainty ($\frac{0.624+0.003}{100}$) remains sizable, although it is understated by 25% comparing to the estimates when switching cost is incorporated. The reason is, as we mentioned previously, that switching cost raises the worker's reservation match and there must be larger variation in the match-level shock in order to explain the observed mobility pattern from data.

Another way to argue that wage risk is underestimated from realized wage changes is to compare the variance of the realized match-specific wage change with the variance of the match-level shock. Figure 7 plots the variance of changes in match quality over the first 30 periods of work life, for a simulated population of 10,000 workers of the same type. The top solid line is the true match-level wage risk estimated by the model. Canonical wage dynamics models attribute the variation of wage changes to uncertainty. When job mobility choices are properly modeled, changes in wages are endogenous and no longer yield correct information on the true wage risk facing workers. The variation of realized match changes (dotted line) underestimates the true wage variation (solid line) which would have taken place without job mobility. A higher switching cost (dashed line) *reduces* the latent wage risk, for shocks are more likely to be reflected in observed wages since workers are less "responsive" to not-so-bad match shocks.

7.3 Decomposing Wage Growth and Inequality for Young Workers

Using the estimated parameters (column (1) of Table 6), I simulate the wage histories for 2500 workers of a given type X_i. I then decompose the mean and the variance of simulated wages over the first 40 periods (13 years) of the life cycle. The primary interest is to evaluate the contribution from the match and the individual component to wages over time. The left panel of Figure 7 examines the experience profile of mean wages. Both the match component and individual component drive wage growth. The growth in $E(u_t)$ is due to the positive experience effect. The growth in $E(a_t)$ is entirely due to job mobility, without which $E(a_t)$ would be declining over time given that the estimated mean return to tenure is negative. Good shocks are preserved and bad shocks could be recovered through job changes. As the extent of job-to-job transitions decreases with experience, the growth of $E(a_t)$ gradually slows down, generating the concave experience profile of wages.

The right panel of Figure 7 decomposes the age profile of wage inequality for young workers. At the beginning of life, almost all of the wage inequality is from variation in individual heterogeneity (i.e. individual's general ability). As worker accumulates labor market experience, the contribution from the worker-firm match quickly rises as a result of the permanent match shocks, job-to-job transitions, and heterogeneous return to tenure. Indeed, after 10 years from the beginning of life, the variation of match-level wages contribute as much as the wage variation at the person level to log wage inequality. In other words, differences in labor market histories are the main driving force behind the increasing inequality over life. This result is similar to recent findings from Huggett, Ventura, and Yaron (2011). Note that transitory shocks and/or measurement errors play a significant role in explaining cross-sectional wage inequality throughout life.

8 Conclusion

This paper jointly estimates a model of wage dynamics and a model of job mobility with switching cost. I consider two sources of wage shocks: shocks at the worker-firm match level and shocks at the individual level which persist across jobs. The key identifying restriction is that match-level shocks affect the worker's job mobility decision, while the person-level shock is independent of job mobility. Given a model of on-the-job search, distinguishing match-level shocks from person-level shocks is economically important. I identify the true wage risk prior to job mobility, which is more than twice as large as the wage risk estimated using observed wages alone. I also show that job mobility is a valuable channel in response to the match-level wage shocks. The extent of latent wage risk and the value of job mobility depend critically on the magnitude of the switching cost and the variation of match-level wage shocks.

There are a few policy implications from this paper. During an economic recession, the toll on displaced workers usually attracts much attention. However, this paper suggests that the welfare of employed workers could also be greatly reduced in recessions. For employed workers, recessions could lead to a rise in job switching cost or a decline in the offer arrival probability. Under either case, job mobility is less valuable in terms of its value in moving away from bad match-level shocks. In addition, decreasing activities of job-to-job transitions implies that match-level shocks would appear more persistent in recessions and thereby making the wage risk identified from observed wages larger. This coincides with empirical evidence from the literature finding that idiosyncratic risk is countercyclical.[44] Lastly, the fact that the accumulated match-specific shocks account for much of the wage inequality for workers in their early careers suggests an important role for policy interventions (like job training programs) aiming to help workers recover from the loss of job-specific skills.

This paper can be extended in a number of directions. First, the paper ignores unemployment and only focuses on job mobility decisions made by young male workers. It is useful to extend the model to include transitions between employment to unemployment. This is likely to imply a larger variation of match-level shocks within jobs. Second, the wage process considered in this paper remains a simple one. The variance of match-level shock is constant over time. In a model where learning is allowed, the variation of surprises to match quality is decreasing with tenure as workers learn more about their match-specific productivity (Jovanovic, 1979). It would be interesting to explore the implication of the model where job is an "experience good". In that case, the option value of current job would be declining with tenure, which would presumably dampen the value of job mobility. Third, jobs considered in the paper differ from each other only in offered match qualities. Modeling transitions across jobs that differ in wage risk, return to tenure, or hours of work is left for future research. Each extension would add another state variable in the model and require a careful specification of the preference structure. Finally, an important avenue for future research is to analyze the relation between job mobility and other channels which workers can rely on in response to labor market risk, and to quantify their relative importance for reacting against different types of shocks.

[44] See Storesletten, Telmer, and Yaron (2004).

APPENDIX

A Proof of Proposition 1

For notational convenience, I omit state variables c_i and k_i, leaving a_{ijt} the only state variable in worker's problem. Therefore, the value function defined here is conditional on worker's type c_i and k_i.

Lemma 1. *$V_t^e(a)$ is monotonically increasing in a, for all t.*

Proof. This can be established through backward induction.

$$V_T^e(a_{ijT}) = a_{ijT}$$

which is an increasing function in a_{ijT} trivially. Now suppose $V_{t+1}^e(a_{ijt+1})$ is increasing in a_{ijt+1}. Proving that $V_t^e(a_{ijt})$ is increasing in a_{ijt} concludes the induction. Now,

$$V_t^e(a_{ijt}) = a_{ijt} + \Gamma(1-\lambda^e)E_t\left[V_{t+1}^e(a_{ijt+1})\right] + \Gamma\lambda^e E_t\left[max\ V_{t+1}^e(a_{ijt+1}), V_{t+1}^e(a_{ij't+1}) - k_i\right] \tag{1}$$

We know that

$$E_t\left[V_{t+1}^e(a_{ijt+1})\right] = \int V_{t+1}^e(a_{ijt+1})dF(a_{ijt+1}|a_{ijt})$$

By the assumptions on the match process, it is easy to show that $F(a_{ijt+1}|a_{ijt}^1)$ first-order stochastically dominates $F(a_{ijt+1}|a_{ijt}^2)$, for any $a_{ijt}^1 > a_{ijt}^2$. This implies that $\int v(k)dF(k|w^1) > \int v(k)dF(k|w^2)$ for any increasing function v. Since by assumption $V_{t+1}^e(a_{ijt+1})$ is increasing in its argument and a_{ijt+1} is increasing in a_{ijt}, $V_{t+1}^e(a_{ijt+1})$ is also increasing in a_{ijt}. Hence we have established that $E_t\left[V_{t+1}^e(a_{ijt+1})\right]$ is increasing in a_{ijt}.

Suppose $V_{t+1}^e(a_{ijt+1}) > V_{t+1}^e(a_{ij't+1}) - k_i$. From equation (1), it is easy to show that $V_t^e(a_{ijt})$ is increasing in its argument. Suppose $V_{t+1}^e(a_{ijt+1}) < V_{t+1}^e(a_{ij't+1}) - k_i$. Since a_{ijt+1} is increasing in a_{ijt}, given k_i, $a_{ij't+1}$ must also be. Then $E_t\left[V_{t+1}^e(a_{ij't+1}) - k_i\right]$ is increasing in a_{ijt} and hence $V_t^e(a_{ijt})$ is increasing in its argument. □

Corollary 1. *If there is no job switching cost, then a worker always chooses a reservation value $r(a_{jt})$ such that $r(a_{jt}) = a_{jt}$.*

Proof. The reservation match value satisfies:

$$V_t^e(a_{ijt}) = V_t^e(r(a_{ijt})) \tag{2}$$

By Lemma 1, we conclude $r(a_{ijt}) = a_{ijt}$ for all t. □

Lemma 2. *The reservation match value $h_t(a,k)$ is monotonically increasing in a, for all t.*

Proof. The reservation match value satisfies:

$$V_t^e(a_{ijt}) = V_t^e(h_t(a_{ijt}, k_i)) - k_i$$

By the implicit function theorem[45],

$$\frac{\partial h_t(a_{ijt}, k_i)}{\partial a_{ijt}} = \frac{V_t^e(a_{ijt})'}{V_t^e(h_t(a_{ijt}, k_i))'} > 0$$

□

Lemma 3. *$V_t^e(a)'$ is monotonically increasing in a, for all t=1... T-1.*

Proof. Let $\Phi(\cdot)$ and $\phi(\cdot)$ denote cumulative and density function of the offer distribution respectively. Differentiating equation 1 w.r.t. a_{ijt}

$$\begin{aligned}\frac{\partial V_t^e(a_{ijt})}{\partial a_{ijt}} &= 1 + \Gamma(1-\lambda^e)E_t(V_{t+1}^e(a_{ijt+1})') \\ &+ \Gamma\lambda^e E_t\left[V_{t+1}^e(a_{ijt+1})'\Phi(h_{t+1}(a_{ijt+1})) + V_{t+1}^e(a_{ijt+1})\phi(h_{t+1}(a_{ijt+1}, k_i))\frac{\partial h_{t+1}(a_{ijt+1}, k_i)}{\partial a_{ijt+1}}\right] \\ &- \Gamma\lambda^e E_t\left[(V_{t+1}^e(h_{t+1}(a_{ijt+1})) - k_i)\phi(h_{t+1}(a_{ijt+1}, k_i))\frac{\partial h_{t+1}(a_{ijt+1}, k_i)}{\partial a_{ijt+1}}\right] \\ &= 1 + \Gamma(1-\lambda^e)E_t(V_{t+1}^e(a_{ijt+1})') + \Gamma\lambda^e E_t\left[V_{t+1}^e(a_{ijt+1})'\Phi(h_{t+1}(a_{ijt+1}))\right] \qquad (3)\end{aligned}$$

where the last step follows because by definition, $V_{t+1}^e(h_{t+1}(a_{ijt+1})) - k_i) = V_{t+1}^e(a_{ijt+1})$. From (3), we obtain

$$\begin{aligned}\frac{\partial^2 V_t^e(a_{ijt})}{\partial a_{ijt}^2} &= \Gamma(1-\lambda^e)E_t(\frac{\partial^2 V_{t+1}^e(a_{ijt+1})}{\partial a_{ijt+1}^2}) \\ &+ \Gamma\lambda^e E_t\left[\frac{\partial^2 V_{t+1}^e(a_{ijt+1})}{\partial a_{ijt+1}^2}\Phi(h_{t+1}(a_{ijt+1}) + V_{t+1}^e(a_{ijt+1})'\phi(h_{t+1}(a_{ijt+1}, k_i))\frac{\partial h_{t+1}(a_{ijt+1}, k_i)}{\partial a_{ijt+1}}\right] \qquad (4)\end{aligned}$$

From Lemma 1 and Lemma 2, we know the last term in equation (4) must be positive. In a backward induction argument, equation (4) essentially proves the core of the induction: if $\frac{\partial^2 V_{t+1}^e(a_{ijt+1})}{\partial a_{ijt+1}^2}$ is positive, then $\frac{\partial^2 V_t^e(a_{ijt})}{\partial a_{ijt}^2}$ must also be positive. Therefore, to complete the proof, we only need to show that the claim is true in the last period. In period T, $\frac{\partial^2 V_T^e(a_{ijT})}{\partial a_{ijT}^2} = 0$. Moving one period backwards,

$$V_t^e(a_{ijT-1}) = a_{ijT-1} + \Gamma(1-\lambda^e)E_{T-1}\left[a_{ijT}\right] + \Gamma\lambda^e E_{T-1}\left[max\ a_{ijT}, a_{ij'T} - k_i\right]$$

It is straightforward to show that $\frac{\partial^2 V_t^e(a_{ijt})}{\partial a_{ijt}^2} = \Gamma\lambda^e E_{T-1}(\phi(a_{ijT} + k_i)) > 0$. □

[45] $V(a)'$ denotes partial derivative of the value function w.r.t a.

So far we have established that the worker's value function is monotonically increasing and convex. We are now ready to derive properties of the reservation wage in the presence of switching cost.

Proposition 1. *If $k \neq 0$, a worker's optimal strategy is to set a reservation match value $h_t(a_{jt}, k)$ where a worker chooses to move if and only if there is an offer such that $a^o_{j'} > h_t(a_{jt}, k)$. Furthermore, for all $t = 1 \dots T-1$, $h_t(a_{jt}, k)$ satisfies the following properties: (1) $h_t(a_{jt}; c, k) > a_{jt}$ if $k > 0$ (2) $\frac{\partial h_t(a_{jt}; c, k)}{\partial k} > 0$ (3) $0 < \frac{\partial h_t(a_{jt}; c, k)}{\partial a_{jt}} < 1$.*

Proof. The reservation match value is defined by:

$$V^e_t(a_{ijt}) = V^e_t(h_t(a_{ijt}, k_i)) - k_i \tag{5}$$

Given $k_i > 0$ and the value function is monotonically increasing, one can easily prove by contradiction that $h_t(a_{ijt}) > r(a_{jt}) = a_{ijt}$.

To show the second property, recall that $V^e_t(a)'$ is positive (Lemma 1) and monotonically increasing (Lemma 3). Therefore,

$$V^e_t(h_t(a_{ijt}, k_i))' > V^e_t(a_{ijt})' > 0$$
$$0 < \frac{\partial h_t(a_{ijt}, k_i)}{\partial a_{ijt}} = \frac{V^e_t(a_{ijt})'}{V^e_t(h_t(a_{ijt}, k_i))'} < 1$$

Next, we show that $h_t(a, k)$ is monotonically increasing in k. Differentiating equation (5) with respect to k, we get

$$\frac{\partial h_t(a_{ijt}, k_i)}{\partial k_i} = \frac{\frac{\partial V^e_t(a_{ijt})}{\partial k_i} + 1}{V^e_t(h_t(a_{ijt}, k_i))'}$$

We know the denominator must be positive from Lemma 1. What remains to be shown is that the nominator is also positive. We prove this by backward induction.

Let us first examine $\frac{\partial V^e_t(a_{ijt})}{\partial k_i}$ in $T-1$. It is easy to show that

$$\frac{\partial V^e_{T-1}(a_{ijT-1})}{\partial k_i} = -\Gamma \lambda^e E_{T-1}(1 - \Phi(a_{ijT} + k_i))$$

whose value lies between (-1,0). Now suppose $\frac{\partial V^e_{t+1}(a_{ijt+1})}{\partial k_i} \in (-1, 0)$. Differentiating V^e_t w.r.t k

$$\begin{aligned}\frac{\partial V^e_t(a_{ijt})}{\partial k_i} &= 1 + \Gamma(1-\lambda^e)E_t(\frac{\partial V^e_{t+1}(a_{ijt+1})}{\partial k_i}) + \Gamma \lambda^e E_t \left[\frac{\partial V^e_{t+1}(a_{ijt+1})}{\partial k_i} \Phi(h_{t+1}(a_{ijt+1}, k_i)) \right] \\ &\quad + \Gamma \lambda^e E_t \int_{h_{t+1}(a_{ijt+1}, k_i)}^{\infty} \frac{\partial V^e_{t+1}(a_0)}{\partial k_i} d\Phi(a_0) - \Gamma \lambda^e E_t \left[1 - \Phi(h_{t+1}(a_{ijt+1}, k_i))\right]\end{aligned} \tag{6}$$

By the induction assumption, it is obvious that $\frac{\partial V_t^e(a_{ijt})}{\partial k_i} < 0$. Then, since we know[46]

$$\int_{h_{t+1}(a_{ijt+1},k_i)}^{\infty} \frac{\partial V_{t+1}^e(a_0)}{\partial k_i} d\Phi(a_0) > \frac{\partial V_{t+1}^e(h_{t+1}(a_{ijt+1},k_i))}{\partial k_i}(1-\Phi(h_{t+1}(a_{ijt+1},k_i)))$$

Equation (6) then simplifies to

$$\frac{\partial V_t^e(a_{ijt})}{\partial k_i} > \Gamma E_t(\frac{\partial V_{t+1}^e(a_{ijt+1})}{\partial k_i}) > -1$$

assuming the discount factor Γ is between (0,1). Therefore, we have shown that for all $t = 1 \dots T-1$,

$$\frac{\partial V_t^e(a_{ijt})}{\partial k_i} + 1 > 0$$

and hence

$$\frac{\partial h_t(a_{ijt},k_i)}{\partial k_i} > 0$$

$\square$

B Approximating the Value Function

For the model with job switching cost, we need to approximate the value function in order to simulate wage and job mobility histories. I choose to specify a terminal value function at time T_0 and solve the model backwards from T_0. The assumption at $t = T_0$ is that job mobility ceases and there are no match-level wage shocks from $T_0 + 1$ until the end of work life T. Solving the model backwards from T is computationally expensive, since the decision period in the model is four months. In addition, other types of shocks (e.g. health shocks) that influence job mobility choices may become increasingly important as young workers age. Also, the distribution of wage shocks may not be the same as the one when workers were young. Therefore, solving the model backwards from the end of life runs the risk of misspecification and misidentifying the parameters of interest. I set $T = 100$ periods (33.3 years) and $T_0 = 40$ periods (13.3 years).[47]

The value function does not have an analytical solution. Recall that the state variables contain $\{c_i, a_{ijt}, X_i\}$. The continuation value (for a person of c_i and X_i employed by firm j) can be approximated by Monte Carlo simulations. For example, the *Emax* function in equation (1) of the Appendix is given

[46] This builds on that $\frac{\partial V_t^e(a_{ijt})}{\partial k_i}$ is increasing in a, which can be proved by induction. It is easy to see that this holds for period $T-1$.

[47] Solving the model for additional periods should not make any difference, since the average periods of experience for the workers in the data is 12 periods, which is well below T_0.

by

$$\frac{1}{D}\sum_{d=1}^{D}\max(V^e_{t+1}(a^d_{ijt+1}), V^e_{t+1}(a^d_{ij't}) - \mathbf{X_i}'\mathbf{\Upsilon}) \tag{7}$$

where $a^d_{ijt+1}, a^d_{ij't}$ are the d_{th} draws from the distribution of a_{ijt+1} given a_{ijt} (equivalent to a draw from the distribution of η_{it}) and $a^o_{ij't}$ respectively. By model assumptions, these random draws are from independent normal distributions. In practice, for each random variable x, I draw n (n=10) equiprobable values of x_i so that $E(x)$ can be approximated by $\frac{1}{n}\sum_{d=1}^{n} x_d$. [48]

The computational burden from solving the value function arises primarily from the continuous and serially correlated state variables c_i and a_{ijt}. c_i represents persistent unobserved heterogeneity. As it is commonly done in the literature, I discretize its distribution and then solve the value function for different values of c_i. The difficulty with a_{ijt} is that to evaluate value function at t, it is necessary to compute the value function for every possible value of a_{ijt+1} which may arise in $t+1$. The number of possible values of a_{ijt+1} grows exponentially with t, making computation quickly infeasible. To circumvent this issue, I use an interpolation method developed in Bound, Stinebrickner, and Waidmann (2009). The method involves two steps. In the first step, I determine the range of possible values of a_{ijt} that could arise from simulations used to approximate the value function and to evaluate the moments in every period $t = 1, \ldots, T_0$. The second step solves the value function backwards. At each time t, the value function is evaluated at N equally spaced grid point a^n_{ijt}. To calculate the value function at each grid point at time t, I need to calculate the value function at $t+1$ for possible values of a_{ijt+1}. These values of a_{ijt+1} will not correspond to the grid points in $t+1$ in general. Each of the possible value functions at a_{ijt+1} is approximated by interpolating between the two value functions associated with two surrounding grid points a^{n-1}_{ijt+1} and a^n_{ijt+1}. I set the number of grid points in each period to 10. Increasing the number of grids does not change the estimated parameters. We know that the value function is monotonic and well behaved. The value function is computed to simulate the job mobility choice, which is a binary variable. These factors place less demand in the interpolation procedure.

C Details of Simulating the Theoretical Moments

For the model without switching cost, the selection equation does not depend on observables X_i, meaning that the simulated moment can be written as $\hat{f}(\theta; \tau_i, p_1, p_2)$. In this case, I simulate S ($S = 5000$) elements of c_i(return to tenure), and S vectors of job offer and worker-firm match shock.[49] Both

[48] x_i is constructed in the following way. Define a vector of equal spaced values on the interval $[0,1]$: $A = \{0, 1/(n-1), 2/(n-1)\ldots 1\}$. For each element A_i in A, define $A_i^{-1} = F^{-1}(A_i)$, where F is the c.d.f of x (normally distributed in our case). Then x_i is equal to $E(x|A_{i-1}^{-1} \le x \le A_i^{-1}) = n\int_{A_{i-1}^{-1}}^{A_i^{-1}} x dF(x)$.

[49] These normally distributed random variables are constructed through the inversion method. That is, first draw a vector of random variables z from a uniform (0,1) distribution. Evaluating the inverse of cumulative normal distribution $F^{-1}(z)$ yields a vector of normally distributed random variables. The uniform draws z are held fixed and independent of model parameters. This guarantees that MSM objective function varies only with respect to changes in parameters of interest.

the job-offer and match-shock vectors have lengths equal to 40, the maximum length of labor-market experience observed for the worker. Given these draws, I obtain S simulated mobility and wage residuals for the first 40 periods as implied from the model. For the model with switching cost, the difference is that the simulated moment is a function of observables X. Therefore, the simulated mobility and wage histories are type specific (i.e. depend on X). To ease computational burden when approximating the value function, the distribution of c_i is discretized into Q ($Q = 25$) points: $\{c_i^q\}_{q=1}^Q$, each carrying the same probability mass. For each type of X and each type of c_i, I simulate S_q ($S_q = 100$) mobility and wage histories since the beginning of work life. Given X, the total number of simulated job histories is equal to $S_q \times Q$.

For both model specifications, the time when researchers begin to observe a worker's wage depends on τ_i. For example, suppose $\tau_i = x_0$ for a given worker, and a subset of S simulated job histories (denote the set by A) satisfy $M_0 = M_1 = ... = M_{x_0} = 0$. For this particular worker, the simulated moment $\hat{f}$ is then

$$\frac{1}{S_{x_0}} \sum_{s \in A} f(\theta; \hat{\nu_s}, x_0, X_i, p_1, p_2) \tag{8}$$

where S_{x_0} is the number of simulated job histories that are in set A.[50]

D Deriving Analytical Forms of the Moments

For the purpose of illustration, I make two simplifying assumptions. First, assume that complete wage and mobility histories are observed from the beginning of life (period 0) up to period T. Second, there is no switching cost. Switching cost parameters Υ can be identified from exclusion restrictions, as illustrated in Section 5.

Using the notation introduced in Section 5.1, let a^o be an offer, which is a random draw from the normal distribution with mean zero and variance $\sigma_{a_0}^2$. Let a_t^l denote the latent match in period t before job mobility decision is made. Consider the unconditional first moment of mobility and wages:

$$E(M_t) = E(E(M_t|a_t^l)) = \lambda^e E_{a_t^l}\left[\Phi(-\frac{a_t^l}{\sigma_{a_0}})\right] \tag{9}$$

$$\begin{aligned} E(r_t) &= E(a_t) + E(u_t) + E(v_t) \\ &= E(a_t|M_t = 1)E(M_t) + E(a_t|M_t = 0)(1 - E(M_t)) + \mu_\delta \times t \\ &= \sigma_{a_0} E_{a_t^l}\left[\lambda(a_t^l)\right] E(M_t) + \left[E(a_t^l)(1 - \lambda^e) + \lambda^e E_{a^o}\left[E(a_t^l|a_t^l > a^o)\right]\right](1 - E(M_t)) + \mu_\delta \times t \end{aligned} \tag{10}$$

where λ is the inverse Mills ratio: $\lambda(a_t^l) = \frac{\phi(\frac{a_t^l}{\sigma_{a_0}})}{1-\Phi(a_t^l/\sigma_{a_0})}$. Φ and ϕ are standard normal c.d.f and density function respectively. a_t^l is a function of complete history of wage draws: $c, a_0, \{\eta_p\}_{p=1}^t, \{a_{j'p}^o\}_{p=1}^{t-1}$. The

[50] This is the crude accept-reject method. I constrain S_{x_0} to be bounded below (minimum of 2% of total number of simulations) to ensure that S_{x_0} is also large.

unconditional density of a_t^l is a function of $\Theta^m = \{\lambda^e, \sigma^2_{a_0}, \sigma^2_\eta, \sigma^2_c, \mu_c, \mu_\delta\}$.

Consider the unconditional covariance between mobility and wage and autocovariance of mobility:

$$cov(r_k, M_t) = E(r_k|M_t = 1)E(M_t) - E(r_k)E(M_t) \tag{11}$$
$$cov(M_k, M_t) = E(M_t|M_k = 1)E(M_k) - E(M_k)E(M_t) \tag{12}$$

where $t \geq k \geq 0$. The new identification restriction imposed from these covariance moments are essentially $E(r_k|M_t = 1)$ and $E(M_t|M_k = 1)$, which can be written as:

$$\begin{aligned} E(r_k|M_t = 1) &= E(a_k|M_t = 1) + E(u_k|M_t = 1) + E(v_k|M_t = 1) \\ &= E(a_k|M_t = 1) + \mu_\delta \times k \\ E(M_t|M_k = 1) = P(M_t = 1|M_k = 1) &= \frac{P(M_t = 1, M_k = 1)}{E(M_k)} \end{aligned}$$

where the conditional density of $a_k|M_t$ depends on the joint distribution of (a_k, a_t), which again depends on history of wage draws and Θ^m. For the first two periods of the model, we are able to derive analytical expression of $E(a_0|M_1 = 1)$ as follows. The unconditional distribution of a_1^l is normally distributed with mean μ_c and variance of $\sigma^2_{a_1^l} = \sigma^2_{a_0} + \sigma^2_\eta + \sigma^2_c$. The joint distribution of a_1^l and a_0 follows a bivariate normal distribution with means μ_c and zero, variances $\sigma^2_{a_1^l}$ and $\sigma^2_{a_0}$, and correlation coefficient $\rho = \frac{\sigma_{a_0}}{\sigma_{a_1^l}}$. Then,

$$E(r_0|M_1 = 1) = \rho\sigma_{a_0} E_{a^o}\left[\lambda(a^o)\right] \tag{13}$$

where $\lambda(a^o) = -\frac{\phi((a^o - \mu_c)/\sigma_{a_1^l})}{\Phi((a^o - \mu_c)/\sigma_{a_1^l})}$.

Each of these moments above (equation (9)-(12)) is a function of the parameters in Θ^m. Then, given Θ^m, the autocovariance of wage can be written as a function of Θ^u:

$$cov(r_{it}, r_{ik}) - cov(a_{it}, a_{ik}) = \sigma^2_{u_0} + tk\sigma^2_\delta + k\sigma^2_\eta, if\ k < t, \tag{14}$$
$$var(r_{it}) - var(a_{it}) = \sigma^2_{u_0} + t^2\sigma^2_\delta + t\sigma^2_\eta + \sigma^2_v \tag{15}$$

where the LHS is the autocovariance of wage that is unexplained by the match component.

References

Abowd, J., and D. Card (1989): "On the covariance structure of earnings and hours changes," *Econometrica*, pp. 411–445.

Abraham, K., and H. Farber (1987): "Job duration, seniority, and earnings," *The American Economic Review*, 77(3), 278–297.

Altonji, J., and L. Segal (1996): "Small-sample bias in GMM estimation of covariance structures," *Journal of Business & Economic Statistics*, 14(3), 353–366.

Altonji, J., and R. Shakotko (1987): "Do wages rise with job seniority?," *The Review of Economic Studies*, 54(3), 437–459.

Altonji, J., A. Smith, and I. Vidangos (2009): "Modeling Earning Dynamics ," *NBER working paper*.

Altonji, J., and N. Williams (2005): "Do Wages Rise with Job Seniority? A Reassessment," *Industrial and Labor Relations Review*, 58(3), 370–397.

Baker, M., and G. Solon (2003): "Earnings dynamics and inequality among Canadian men, 1976-1992: Evidence from longitudinal income tax records," *Journal of Labor Economics*, 21(2), 289–321.

Bils, M., Y. Chang, S. Kim, and H. Hall (2009): "Heterogeneity and Cyclical Unemployment," *NBER Working Paper*.

Blundell, R., and L. Pistaferri (2003): "Income Volatility and Household Consumption: The Impact of Food Assistance Programs," *The Journal of Human Resources*, 38, 1032–1050.

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption Inequality and Partial Insurance," *American Economic Review*, 98(5), 1887–1921.

Bound, J., T. Stinebrickner, and T. Waidmann (2009): "Health, economic resources and the work decisions of older men," *Journal of Econometrics, forthcoming*.

Burdett, K. (1978): "A theory of employee job search and quit rates," *The American Economic Review*, pp. 212–220.

Carroll, C. (1992): "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity*, 1992(2), 61–156.

Cunha, F., J. Heckman, and S. Navarro (2005): "Separating uncertainty from heterogeneity in life cycle earnings," *Oxford Economic Papers*, 57(2), 191.

Deaton, A. (1992): *Understanding Consumption*. Oxford University Press.

Dey, M., and C. Flinn (2005): "An Equilibrium Model of Health Insurance Provision and Wage Determination," *Econometrica*, 73(2), 571–627.

Flinn, C. (1986): "Wages and job mobility of young workers," *The Journal of Political Economy*, pp. 88–110.

Flinn, C., and J. Heckman (1982): "New methods for analysing structural models of labour force dynamics," *Journal of Econometrics*, 18, 115–168.

Gottschalk, P. (2005): "Downward nominal-wage flexibility: real or measurement error?," *Review of Economics and Statistics*, 87(3), 556–568.

Gottschalk, P., and R. Moffitt (1994): "The Growth of Earnings Instability in the US Labor Market," *Brookings Papers on Economic Activity*, 1994(2), 217–272.

Gourinchas, P., and J. Parker (2002): "Consumption Over the Life Cycle," *Econometrica*, 70(1), 47–89.

Gruber, J. (1997): "The consumption smoothing benefits of unemployment insurance," *The American Economic Review*, pp. 192–205.

Guvenen, F. (2007): "Learning your earning: Are labor income shocks really very persistent?," *American Economic Review*, 97(3), 687–712.

Haider, S. (2001): "Earnings instability and earnings inequality of males in the United States: 1967-1991," *Journal of Labor Economics*, 19(4), 799–836.

Heathcote, J., K. Storesletten, and G. Violante (2009): "Quantitative Macroeconomics with Heterogeneous Households," *working paper*.

Heckman, J. (1981): "The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process," *Structural analysis of discrete data with econometric applications*, pp. 179–195.

Hey, J., and C. McKenna (1979): "To move or not to move?," *Economica*, 46(182), 175–185.

Hornstein, A., P. Krusell, and G. Violante (2011): "Frictional wage dispersion in search models: A quantitative assessment," *American Economic Review*, 101(7), 2873–98.

Huggett, M., G. Ventura, and A. Yaron (2011): "Sources of Lifetime Inequality," *The American Economic Review*, 101(7), 2923–54.

Jensen, S., and S. Shore (2008): "Changes in the Distribution of Income Volatility," *working paper*.

Jovanovic, B. (1979): "Job Matching and the Theory of Turnover," *The Journal of Political Economy*, 87(5), 972.

Kaplan, G. (2009): "Moving back home: Insurance against labor market risk," *working paper, NYU*.

Low, H. (2005): "Self-insurance in a life-cycle model of labour supply and savings," *Review of Economic Dynamics*, 8(4), 945–975.

Low, H., C. Meghir, and L. Pistaferri (2010): "Wage Risk and Employment Risk over the Life Cycle," *The American Economic Review*, 100(4), 1432–1467.

Low, H., and L. Pistaferri (2010): "Disability Risk, Disability Insurance and Life Cycle Behavior," *working paper*.

Meghir, C., and L. Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1), 1–32.

——— (2011): "Earnings, Consumption and Life Cycle Choices," *Handbook of Labor Economics*, 4, 773–854.

Moffitt, R., and P. Gottschalk (2011): "Trends in the Transitory Variance of Male Earnings in the U.S, 1970-2004," *NBER working paper*.

Mortensen, D. (1986): "Job Search and Labor Market Analysis," *Handbook of Labor Economics*, 2, 849–919.

Nagypal, E. (2005): "On the extent of job-to-job transitions," *working paper*.

Pischke, J. (1995): "Measurement error and earnings dynamics: Some estimates from the PSID validation study," *Journal of Business & Economic Statistics*, 13(3), 305–314.

Postel-Vinay, F., and J. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70(6), 2295–2350.

Rogerson, R., R. Shimer, and R. Wright (2005): "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, 43(4), 959–988.

Storesletten, K., C. Telmer, and A. Yaron (2004): "Cyclical dynamics in idiosyncratic labor market risk," *Journal of Political Economy*, 112(3), 695–717.

Topel, R. (1991): "Specific capital, mobility, and wages: Wages rise with job seniority," *Journal of Political Economy*, pp. 145–176.

Topel, R., and M. Ward (1992): "Job mobility and the careers of young men," *The Quarterly Journal of Economics*, pp. 439–479.

Van Der Berg, G. (1992): "A structural dynamic analysis of job turnover and the costs associated with moving to another job," *The Economic Journal*, 102(414), 1116–1133.

Wolpin, K. (1992): "The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth," *The Journal of Political Economy*, 100(3), 535–560.

Table 1: Summary Statistics, SIPP 1996

Variable	Mean	Standard Deviation
Demographics		
Age	26.45	2.78
White	0.75	0.43
Some college or more	0.57	0.49
Metropolitan	0.82	0.38
Own a house	0.49	0.50
Married	0.50	0.50
Labor market variables		
Wages	11.47	5.39
Hours of work per week	43.26	8.61
Proportion of job-job transition	0.10	0.30
Durations of jobs (exclu. right-censored)	6.96	6.75
Elapsed job duration in the first observation period	7.74	7.37
Total number of observations	9688	

Note: Wages are deflated using monthly CPI-Urban (CPI=1 in 1996:1) and averaged over a four-month period (per wave).

Table 2: Total Number of Job Changes (in percentages) and Within-job Wage Growth, by Experience

	Number of job changes					Within-job Δw	
	0	1	2	3	4+	Mean	s.e
Quartiles of initial labor market experience (period)							
Less than 25th (0-2)	31.2	29.3	21.5	10.9	7.1	0.031	0.005
25-50 (3-6)	46.8	29.9	15.3	6.1	1.9	0.020	0.005
50-75 (7-13)	60.0	23.5	11.2	4.2	1.1	0.020	0.006
More than 75th (>13)	72.1	18.9	7.6	1.0	0.3	0.013	0.005
Total	52.2	25.5	14.0	5.6	2.6	0.021	0.003

Table 3: Tobit Regression of Total Number of Job Changes

Married	-0.11
	(0.12)
Own a house	0.55***
	(0.20)
White	0.08
	(0.14)
High school	-0.06
	(0.22)
At least some college	-0.21
	(0.21)
Initial experience level	-0.10***
	(0.01)
Constant	0.93***
	(0.21)
Observations	1211

Note: Tobit regression of total number of job changes on observed personal characteristics. Regressors include a constant, labor market experience in the first observation period, and dummy variables of marital status, whether individual owns a house, race and education. Asymptotic standard errors are in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4: Dynamics of Wage and Job Mobility

	SIPP Sample		Simulated Sample	
	(1a)	(2a)	(1b)	(2b)
Variable	$ln(w_t)$	M_t	$ln(w_t)$	M_t
$ln(w_{t-1})$	0.554	-0.091	0.502	-0.116
	(0.011)	(0.000)	(0.002)	(0.003)
$ln(w_{t-2})$	0.328	-0.018	0.392	-0.015
	(0.011)	(0.217)	(0.002)	(0.003)
M_t	0.010	-	-0.001	-
	(0.009)		(0.002)	
M_{t-1}	0.037	0.127	-0.010	0.079
	(0.009)	(0.000)	(0.002)	(0.003)
M_{t-2}	-0.016	0.037	-0.020	0.067
	(0.009)	(0.002)	(0.002)	(0.003)
R^2	0.70	0.04	0.73	0.06
Observations	7266	7266	145320	145320

Note: OLS regressions of log wage and job mobility. Regressors include a constant, one- and two-period lagged log wage and mobility. Standard errors (in parenthesis) are clustered by person.

Table 5: Probit Regressions of Job Mobility on Past Wage Changes

	(1)	(2)	(3)	(4)	(5)
Wthin job wage growth in$t-1$	-0.202	-0.256	-0.497	-0.421	-0.601
	(0.088)	(0.093)	(0.104)	(0.132)	(0.145)
Within job wage growth in $t-2$	-	-	-	-0.176	-0.472
				(0.104)	(0.126)
Job tenure in $t-1$	-	-0.036	-0.025	-	-0.029
		(0.003)	(0.005)		(0.007)
$ln(w_{t-2})$	-	-	-0.470	-	-
			(0.062)		
$ln(w_{t-3})$	-	-	-	-	-0.372
					(0.074)
Experience in $t-1$	-	-	-0.006	-	0.001
			(0.005)		(0.007)
Observations	6381	6381	6381	4766	4766

Note: Probit regressions of job mobility on lagged within-job wage growth. The dependent variable is a job change indicator (M=1 if job change occurs). Regressors include a constant, one- and two-period lagged within-job wage growth and wage levels, completed job tenure, work experience, and two- and three-period lagged level of wage. Standard errors (in parenthesis) are clustered by person.

Table 6: Estimated Model Parameters

	With Job-switching Cost	No Job-switching Cost	No Match-level Shocks	No Match-specific Wages
	(1)	(2)	(3)	(4)
Wage shocks				
$\sigma_\eta^2 \times 100$	0.986	0.624	-	-
	(0.007)	(0.020)		
$\sigma_\zeta^2 \times 100$	0.005	0.003	0.176	0.321
	(0.000)	(0.001)	(0.033)	(0.001)
σ_v^2	0.032	0.037	0.036	0.046
	(0.000)	(0.001)	(0.003)	(0.002)
Random growth factor				
μ_c	-0.008	-0.006	-	-
	(0.001)	(0.001)		
$\sigma_c^2 \times 10000$	0.110	0.496	-	-
	(0.012)	(0.226)		
μ_δ	0.011	0.004	-0.002	-
	(0.001)	(0.001)	(0.002)	
$\sigma_\delta^2 \times 10000$	0.274	0.010	0.000	0.000
	(0.003)	(0.008)	(0.004)	(0.004)
Heterogeneity				
$\sigma_{a_0}^2$	0.003	0.015	0.065	-
	(0.000)	(0.001)	(0.010)	
$\sigma_{u_0}^2$	0.066	0.084	0.072	0.057
	(0.000)	(0.002)	(0.005)	(0.008)
Other labor market parameters				
λ^e	0.388	0.849	0.730	-
	(0.001)	(0.053)	(0.022)	
$\gamma^{constant}$	0.399	-	-	-
	(0.017)			
$\gamma^{married}$	-0.377	-	-	-
	(0.008)			
γ^{house}	-0.294	-	-	-
	(0.009)			
γ^{hs}	0.092	-	-	-
	(0.029)			
$\gamma^{college}$	-0.473	-	-	-
	(0.033)			
γ^{white}	-0.298	-	-	-
	(0.011)			

Note: Bootstrap standard errors are in parentheses. σ_η^2, σ_ζ^2 and σ_v^2 are, respectively, the variances of match- and person-level shock, and transitory shock (measurement error). c and δ are the random growth factor at match and person level respectively. $\sigma_{a_0}^2$ is the heterogeneity in the match values of job offers. $\sigma_{u_0}^2$ is the heterogeneity in the person-component of wages at the start of work life. λ^e is the offer arrival probability. γ's are switching cost parameters.

Table 7: Value of Job Mobility in Response to Match-level Wage Shocks

A. Value of Job Mobility for Different Switching Costs

	Switching Cost	ΔV_0	$\Delta V_0/V_0$
Single, do not own a house, college graduate, white	-0.37	0.84	0.13
Married, do not own a house, high school graduate, white	-0.18	0.81	0.14
Single, do not own a house, high school graduate, white	0.19	0.64	0.14
Single, do not own a house, less than high school, black	0.40	0.58	0.14
Single, do not own a house, high school graduate, black	0.49	0.56	0.14

B. Value of Job Mobility for Different Variations of Match-level Shocks

	$0.5 \times \sigma_\zeta^2$	σ_ζ^2	$2 \times \sigma_\zeta^2$
ΔV_0	0.40	0.64	0.93

Figure 1: Decision Process of Job Mobility

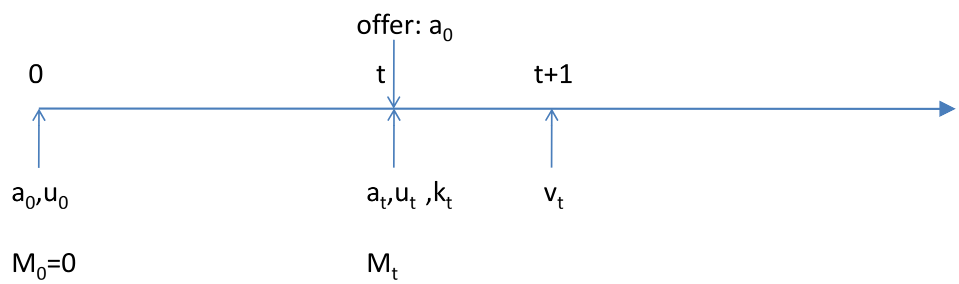


Figure 2: Match-specific Wages and Job Mobility

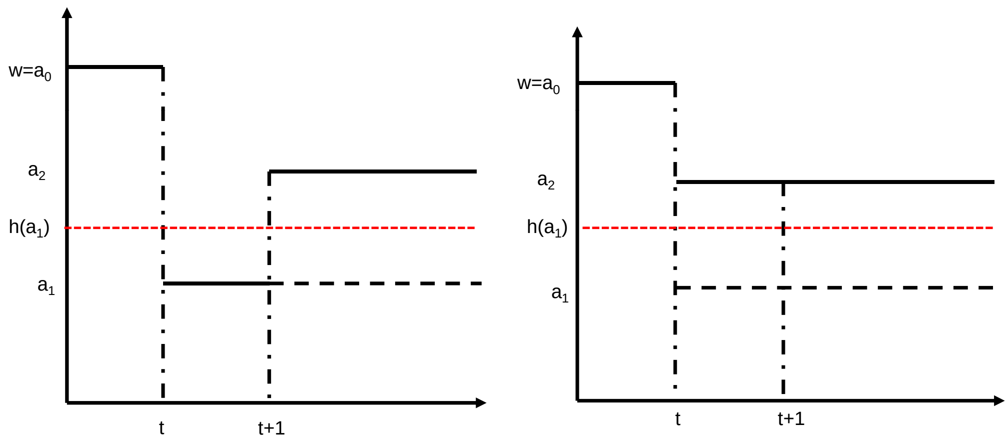


Figure 3: Distributions of Within- and Between-job Wage Growth

Figure 3A: Distribution of within and between real wage growth

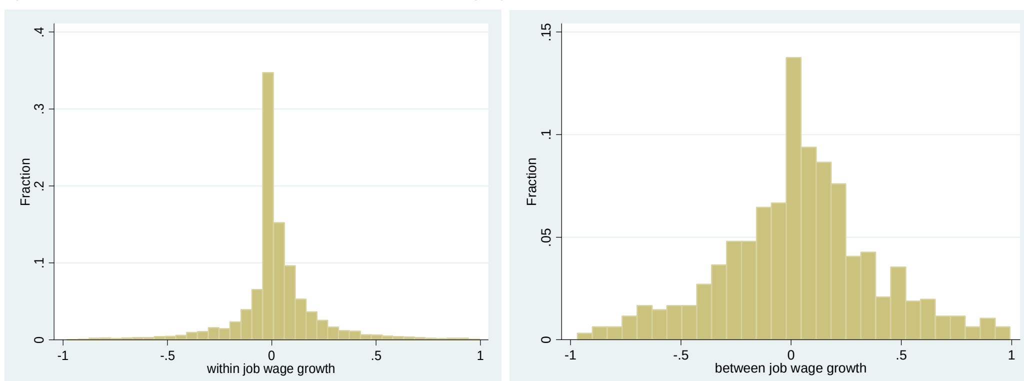


Figure 3B: Distribution of within and between nominal wage growth

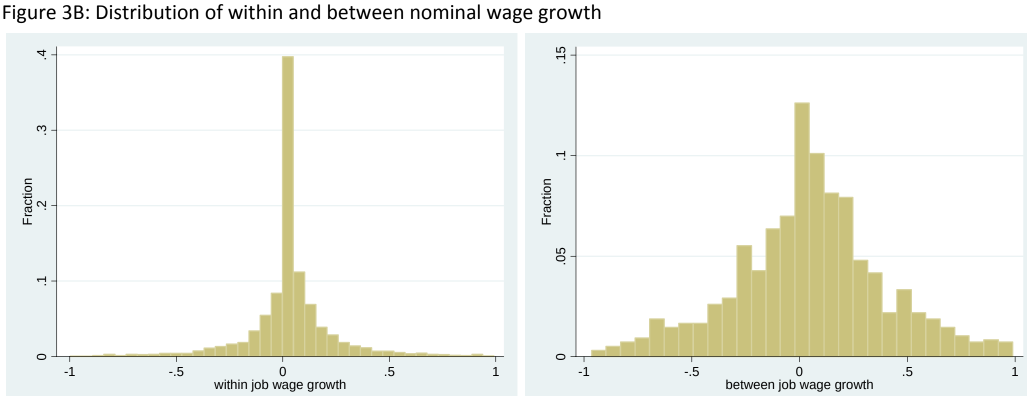
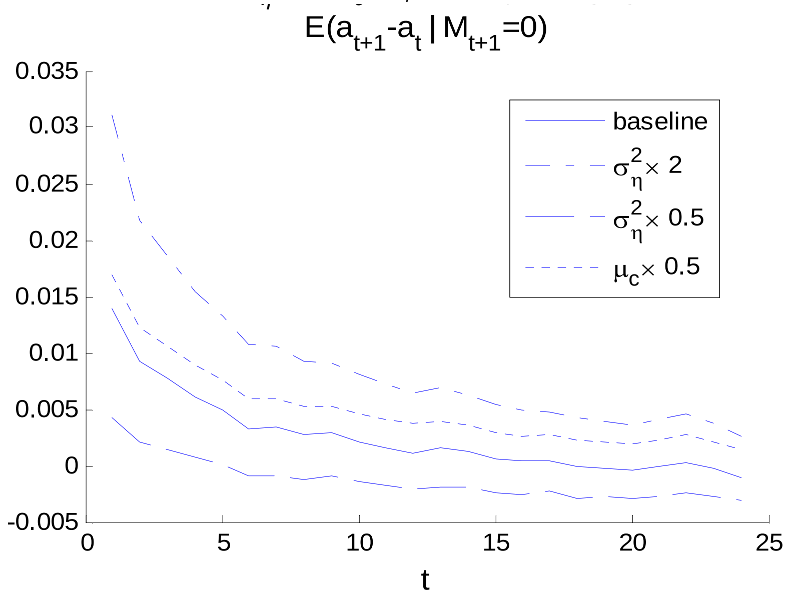
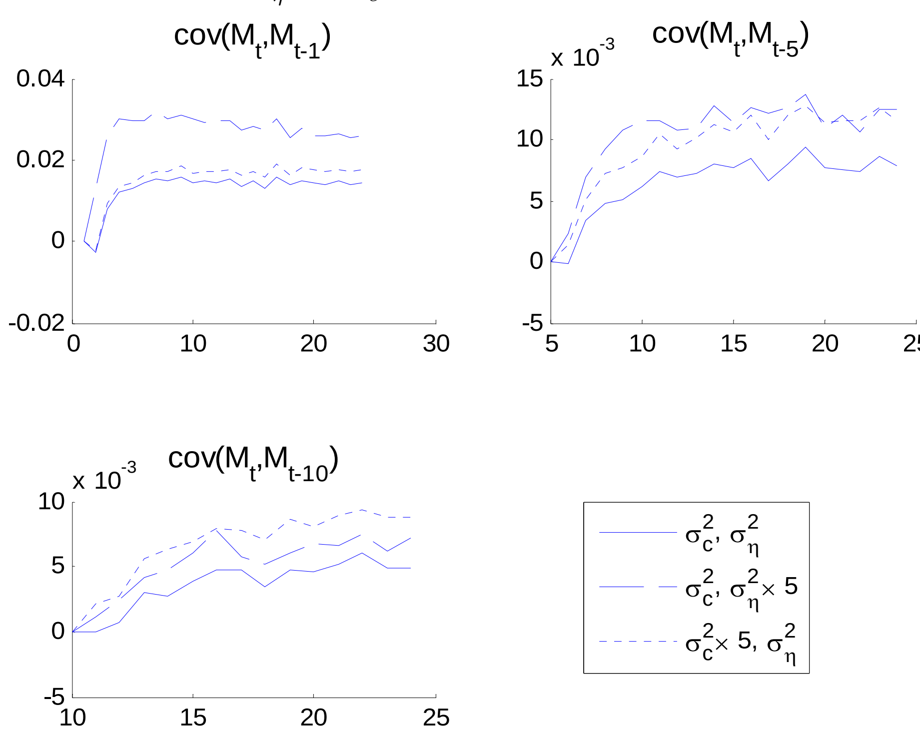


Figure 4: The Effect of σ_η^2 and μ_c on Within-job Wage Growth Over Time



Note: Simulated from the model without switching cost. Baseline: $u_c = -0.006$, $\sigma_\eta^2 = 0.0062$

Figure 5: The Effect of σ_η^2 and σ_c^2 on the Autocovariance of Job Mobility Over Time



Note: Simulated from the model without switching cost. Baseline: $\sigma_c^2 = 4.96 \times 10^{-4}$, $\sigma_\eta^2 = 0.0062$

Figure 6: Changes in the Variance of Realized Match Value Over Time, for a Given Type of Worker

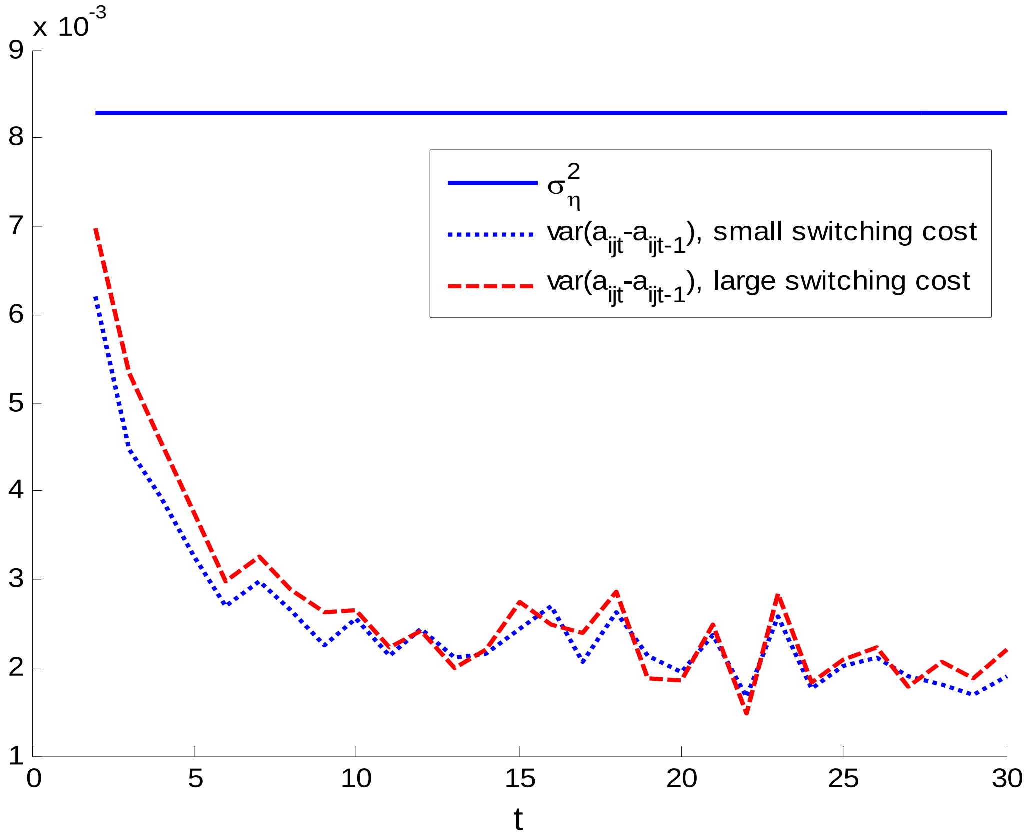
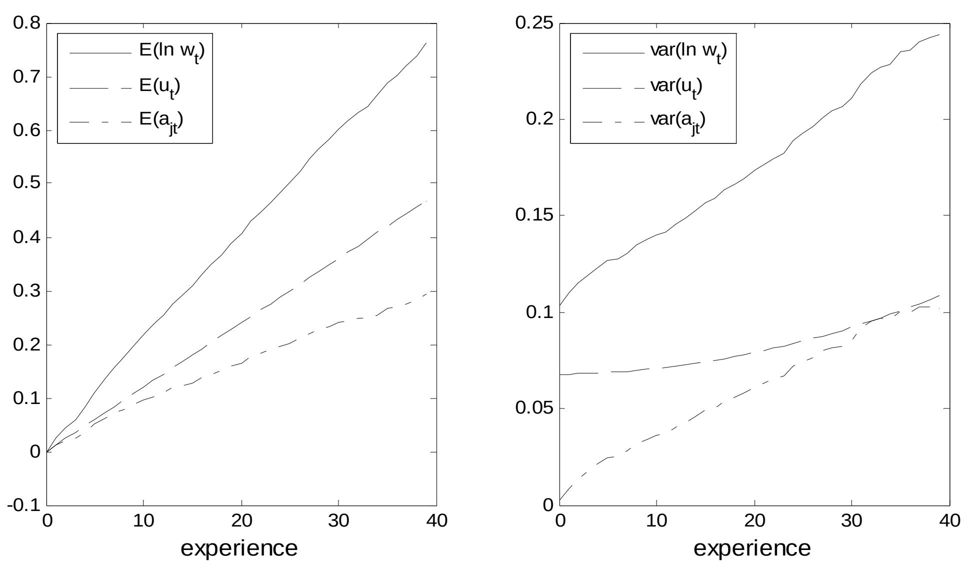


Figure 7: Decomposing the Experience-Profile of Log Wages and Variances of Log Wages



Note: The left panel decomposes the contribution of individual- and match-component to the mean of log wage residual. The right panel decomposes the contribution of individual- and match-component to the variance of log wage residual.