Wage Risk, On-the-job Search and Partial Insurance

Kai Liu
Johns Hopkins University *

JOB MARKET PAPER
This version: November 2010

Abstract

This paper shows that job mobility is a valuable insurance channel against labor market risk for employed workers. I construct a model of wage dynamics jointly with a structural dynamic model of job mobility with job-switching cost. The key feature of the model is the specification of wage shocks at the worker-firm match level, for these shocks can be insured against by job mobility. In addition to allowing job mobility to emerge as a valuable channel for employed workers to self-insure against such match-level wage shocks, the model enables the econometrician to identify the true wage risk prior to job mobility, which will generally be larger than wage variation observed after job mobility has taken place. The model is estimated using panel data of wage and job mobility from the Survey of Income and Program Participation. The first result is that the variance of match-level shocks is large, and the consequent insurance value of job mobility is nearly a quarter of lifetime expected utility. It is more valuable for workers whose initial match draws and/or switching cost is small. The second result is that true wage risk is twice as large as the wage variance observed after job mobility, which is what other papers in the literature have called wage risk. This suggests a very different picture of the risks and welfare for employed workers in the labor market.

JEL: D91, J31, J62

*Email: kliu@jhu.edu. I am grateful to my advisors, Robert Moffitt and Stephen Shore, for guidance and support. I wish to thank Peter Gottschalk and Eswar Prasad for detailed comments and encouragement. I also would like to thank Joe Altonji, Pedro Carneiro, Luigi Pistaferri, Gary and Becky Shaw, Richard Rogerson, Alfonso Rosolia, Kevin Thom, Timen Woutersen, Frank Vella, seminar participants at Johns Hopkins, conference participants at the World Congress of the Econometric Society in August 2010, Society for Economic Dynamics Annual Meeting in July 2010, MOOD Workshop in June 2010, Society of Labor Economists Annual Meeting in June 2010 and QSPS Workshop in May 2010. All remaining errors are mine.
1 Introduction

Understanding how much idiosyncratic risk people face and their channels of insurance against risk is an important research agenda. 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 correctly identified wage risk: both on the level of risk and the persistence of wage innovations.

In most papers, wage risk is identified from the variance of decomposed wage residuals in a panel data model (called “error component model”). Variation and persistence of shocks are revealed by examining the autocovariances of observed wage outcomes. Implicitly in these models, wages and changes in wages are assumed exogenous. However, if we consider employment choices workers make, certain components of observed wages and changes of wages are endogenous. In a frictional labor market, people could select the level of their wages through job mobility. Moreover, workers may be able to self-insure against certain shocks at firm level by switching employers, so both the degree of wage risk and the persistence of shocks could be misidentified in error component models. With shocks mixed with endogenous choice and self-insurance behavior, it is difficult to assess the true welfare cost of wage risk, derive empirical implications of precautionary behavior and evaluate the consequences of government policy interventions. To understand the true wage risk, it is essential to specify sources of shocks and model individual's economic choice together with the wage process\(^2\).

This paper shows that job mobility is a valuable insurance channel against labor market risk for employed workers. I build and estimate an error component model of 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. Shocks to switching cost introduce non-wage risk for employed workers. I distinguish two sources of wage shocks: shocks at worker-firm match level and shocks at individual level which apply to all firms and matches. The main contributions of this approach are twofold. First, I show that worker’s job-job transition serves as a valuable channel for employed workers to self-insure against match-level wage shocks and

---


\(^2\) Recent papers by Low, Meghir, and Pistaferri (2010) and Alkonji, Smith, and Vidangos (2009) also make important contributions in this direction. I illustrate the differences between this paper and their papers in the next section.
non-wage shocks. Second, the model is capable of recovering true wage risk that workers experience prior to job mobility. Observed wage changes are after job mobility decisions and hence are different from the true shocks that occur prior to job mobility.

Job-job transition is a prominent phenomenon in the US labor market. It takes place more than twice as often as transition from employment to unemployment\(^3\). While job mobility has been recognized as the primary drive to wage growth particularly for young workers (e.g. Topel and Ward (1992)), little is known about its value of insurance. In this paper, job mobility arises as a channel to insure against match-level shocks, for two reasons. One is that there is friction in the labor market where workers conduct on-the-job search. The other reason, which is unique to this paper, is the distinction between match-level and person-level wage shocks. If all of the shocks are indeed person specific which would carry onto any job, job mobility would never become a channel of insurance against wage shocks.

I decompose 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\(^4\). The match component and individual component follow parallel stochastic processes: each of them evolves from a permanent shock and a random growth factor. The variances of these permanent shocks represent match- and person-level wage risk respectively. The random growth factor in the individual component represents heterogeneous return to experience, and in the match component, it represents heterogeneous return to tenure.

The on-the-job search model characterizes worker’s job mobility behavior given the specified wage process. Workers are heterogeneous in their observed personal characteristics, growth profiles of wages, job switching costs, initial match and individual components of wages sampled in the first period of work life. When a job offer arrives, worker decides whether to stay with the current job (possibly with a new wage rate) or to take the offer and switch jobs. Under the model, job mobility decision can be characterized by a reservation wage which depends on the match component of wage and switching cost. Therefore, the individual wage component is exogenous and independent of job mobility and the match component is endogenous. Furthermore, match-level shock and hence within-job wage change is

\(^3\)See Nagypal (2008). My own calculation leads to the same conclusion.

\(^4\)Empirically it is infeasible to distinguish pure firm effect from pure worker-firm match effect without employer-employee matched data.
also endogenous. Shock to match component persists within a job, which affects job mobility choices in the current and future periods. Empirically, there is strong negative correlation between job mobility and lagged within-job wage change, controlling for observables. Conditional on current match quality, bad match-level shock is more likely to be latent and less persistent when worker is able to locate a higher-match job quickly. Observed wage changes alone provide a misleading picture of the true magnitude and persistence of the risk.

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 match component from person component in the wage residuals, it hinges on the model’s implication that the match component is correlated with job mobility choices but the person component is not. The key findings are the following: (1) Wage risk at match level accounts for the majority of the wage risk facing workers. (2) True wage risk, identified jointly from wage outcomes and mobility choices, doubles the wage risk that is estimated using wage information alone. (3) The insurance value of on-the-job search is nearly a quarter of the life-time expected utility. It is more valuable for workers whose initial match draws and/or switching cost is low. (4) 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 explains a major portion of the variance of wages at the beginning of work life. Over time, the contribution from match component of wages rises and becomes a dominating contributor to the variance of wages. (6) The estimated mean return to tenure is negative and the mean return to experience is positive. There is strong evidence for heterogeneity in 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. It then illustrates the model’s implications to wage risk and to partial insurance. Section 4 introduces the data set to be used and presents empirical evidence from data which motivates this study. Section 5 discusses estimation and identification strategy. Section 6 presents estimation results. Section 7 defines and quantifies latent wage risk and the insurance value of on-the-job search. It also discusses model’s implications to wage growth and inequality. Section 8 concludes, followed by a discussion on 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 low order MA or ARMA process for the transitory component\(^5\). 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.

A more recent line of work attempts to identify the sources of transitory and permanent income shock. 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 shock 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-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.

Both Altonji, Smith, and Vidangos (2009) and Low, Meghir, and Pistaferri (2010), however, assume that worker-firm match does not vary until the match is dissolved\(^6\). 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 this 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. To quantify the true wage risk.

---


\(^6\)Low, Meghir, and Pistaferri (2010) experiment a random walk process of match component with 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.
facing workers and the value of insurance through job mobility, it’s important to separate wage shocks that are match specific from wage shocks that are person specific, the former of which influence worker’s job mobility choice. Estimation result indicates 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. In addition, 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, specification of such model may suffer from arbitrariness without the guidance of a theory of individual choice. In terms of estimation strategy, Low, Meghir, and Pistaferri (2010) identify their wage process building upon a rather ad-hoc 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 analyzing individuals’ channel of insurance beyond risk-free savings in an incomplete market. Empirical work by Blundell, Pistaferri, and Preston (2008) finds that households are able to partially insure against wage shocks beyond savings. This paper connects to this area of research by showing that job mobility through on-the-job search is another useful channel to self-insure against particular types of wage risk.

This paper is also related to the partial equilibrium on-the-job search model in the literature. 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-job transitions observed in the US labor market, where under plausible parameters, the basic on-the-job search model has failed to achieve. It also provides a parsimonious way to explain wage cuts in job-job transitions. Job mobility with wage cut has been

---

7For instance, they include other events such as health shock and labor supply shock in the earning process besides job mobility.
8Examples 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). I should note that this literature primarily focus on channels of consumption insurance. This paper emphasizes job mobility as a channel for self-insuring income, since I do not model consumption-saving choices in the on-the-job search model.
9Burdett (1978) is the pioneer work. See Mortensen (1986); Rogerson, Shimer, and Wright (2005) for a review.
10See Nagypal (2005) and Hornstein, Krusell, and Violante (2007). Several extensions of the model have been proposed. Nagypal (2005) incorporates a Brownian motion process that causes the value of a job to decrease at times, and Nagypal (2008) considers stochastic unemployment risk over a job tenure.
difficult to reconcile in standard search model because, with stationary wage policy, worker chooses to switch jobs only if there exists a job offering higher wage\textsuperscript{11}. In this paper, between-job wage cut could arise out of four situations: a cut in person-component of wage, a negative latent match-level shock, a large positive shock to switching cost and measurement error. Finally, to the best of my knowledge, few papers in the job search literature consider the dynamic impact of job switching cost in determining worker’s job mobility behavior\textsuperscript{12}. The estimated parameters show that switching cost is an important factor influencing the on-the-job search behavior of young workers.

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:

\begin{align}
\ln \bar{\omega}_{ijt} &= \ln \omega_{ijt} + v_{it} \quad (1) \\
\ln \omega_{ijt} &= Z_i' \beta + a_{ijt} + u_{it} \quad (2) \\
a_{ijt+1} &= \begin{cases} 
    a_{ijt+1}^l, & \text{if no job change between } t \text{ and } t + 1 \\
    a_{ij't+1}^o, & \text{if there is job change between } t \text{ and } t + 1
\end{cases} \\
a_{ijt+1}^l = a_{ijt} + c_t + \eta_{ijt+1}, \quad a_{ij't+1}^o \sim N(0, \sigma_{a0}^2) \quad (3) \\
u_{it+1} = u_{it} + \delta_i + \zeta_{it+1} \quad (4)
\end{align}

\textsuperscript{11}Postel-Vinay and Robin (2002); 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.

\textsuperscript{12}Hey and McKenna (1979) and Van Der Berg (1992) are two exceptions.
Assume that

\[
E(\zeta_{it}) = 0, \quad \text{var}(\zeta_{it}) = \sigma_{\zeta}^2 \tag{5}
\]

\[
E(\delta_i) = \mu_\delta, \quad \text{var}(\delta_i) = \sigma_{\delta}^2 \tag{6}
\]

\[
E(v_{it}) = 0, \quad \text{var}(v_{it}) = \sigma_v^2 \tag{7}
\]

\[
\eta_{ijt} \sim N(0, \sigma_\eta^2), c_i \sim N(\mu_c, \sigma_c^2) \tag{8}
\]

with orthogonality between all five of these variables. \(\tilde{w}_{ijt}\) is the observed real log hourly wage for worker \(i\) employed by firm \(j\) in labor age \(t\). \(Z_i\) is a \(K \times 1\) vector of exogenous regressors of observed heterogeneity including a constant. \(\beta\) is a \(K \times 1\) parameter vector. For an employed worker, unexplained 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 transitory shock \(v_{it}\). The former two components evolve independently under two parallel stochastic processes.

Individual component evolves over the life-cycle from an identically and independently distributed permanent random shock \(\zeta_{it}\) and a random growth factor \(\delta_i\) with mean \(\mu_\delta\). It corresponds to the concept of permanent wage in the literature, which is usually thought of as representing return to skill or human capital. The random growth factor \(\delta_i\) has cross-sectional variance \(\sigma_{\delta}^2\). 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, prior to selection between jobs, match component is a random walk process. Let \(a_{ijt+1}\) be the latent match at \(t+1\) prior to job mobility ("L" represents latent). It evolves from random growth factor (drift) \(c_i\) with mean \(\mu_c\) and cross-sectional variance \(\sigma_c^2\) and permanent shock \(\eta_{ijt}\). Assume that \(\eta_{ijt}\) are identically and independently distributed across firm, worker and time. Unlike person-level shocks, I impose distributional assumption on the match-level shocks and random growth factor. One interpretation of the match component is that it is an idiosyncratic firm effect which is a complement to the individual productivity. From the perspective of human capital theory, match component can also be regarded as job-specific human capital. Random

---

13The age profile of wages will be captured through estimated mean of random growth factors \(\delta_i\) and \(c_i\). In the literature, \(Z\) typically includes a constant and a term linear in age and thus \(\mu_\delta\) is normalized to 0. This paper deviates from it, because with job selection, wages are endogenous except in the first period. See Section 5 for more discussions.

14Because the distribution of \(\eta_{ijt}\) is independent of firm, for simplicity, henceforth I drop the \(j\) subscript on \(\eta\).
growth factor $c_i$ measures 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. Shock to 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 firm does not provide enough training to enhance worker’s firm-specific skills (negative $\eta_{it}$), or it adopts a new technology that is complementary to worker’s productivity (positive $\eta_{it}$). In general it consists of both pure match-specific shock and pure firm-specific shock, although without firm level data, distinguishing between these is not feasible. More broadly, match component can be interpreted as any factor that affects worker’s productivity with current firm but not after he leaves for other firms. Random growth factor and permanent shocks to the match component are accumulating only over the current job tenure and will “vanish” after a job change. Individual component measures worker’s general productivity regardless of his employer.

Flinn (1986) and Topel and Ward (1992) show the importance of 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, match level shock 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. Worker’s wage is determined in a simple partial equilibrium on-the-job search model.

Job offer with match specific wage $a^o$ (“o” stands for “offer”) is a random draw from a stationary offer distribution. Assume it follows a normal distribution with mean zero and variance $\sigma_{ao}^2$. Because of the growth profile in the person-component of wage (due to $\delta_i$), the offered levels of wages would

---

15 This is more general than 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). This would, however, introduce complexity to the search model because with an exogenous wage process I need to consider wage offers varying over both levels and growth rate. I leave it for future research.

16 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.

17 There is a distinction between job mobility between and within occupations, as estimates from Kambourov and Manovskii (2009) show return to occupational tenure is large, even after controlling for return to tenure and to experience. I ignore this issue for now. Accounting for mobility across occupation requires another model of selection process. A more interesting case is where variance of match level shock differs among occupations/sectors, so that people initially self-select into these occupations. This is left for future research.
be mean-shifting with worker’s labor market experience\textsuperscript{18}. Offered matches are assumed uncorrelated with worker’s individual wage component, and hence each firm has 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 job mobility decision, worker is perfectly informed of his general productivity \(u_{it}\), match-specific productivity \(a_{ijt}'\) if he chooses to stay and the value of the offer \(a_{ijt}^o\)\textsuperscript{19}. At any time, workers have perfect information about their current match value, expectation of future match values and the distribution of match component in the labor market, but information of 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\textsuperscript{20}.

Transitory shocks are identically and independently distributed across individual and time. Transitory shock absorbs wage shock of no persistence, at either worker-firm match level or at person level. It also includes classical measurement errors on reported wages. I assume that transitory shocks affect wages after 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 operate at both person and match level\textsuperscript{21}. In this case, both permanent and transitory match-level shocks enter into worker’s information set prior to job mobility decisions. Besides increasing computation burdens, it is difficult to justify a homogeneous correlation restriction on the transitory component of wages applying to all workers’ information set. 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-job transitions

\textsuperscript{18}A more difficult case is to allow offered match to depend on current and lagged match values, which is left for future research.

\textsuperscript{19}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).

\textsuperscript{20}This excludes the possibility that parts of these random shocks may be known to workers in advance. See Cunha, Heckman, and Navarro (2005).

\textsuperscript{21}Many papers, e.g. Moffitt and Gottschalk (1995, 2008); Haider (2001); Meghir and Pistaferri (2004), show that there is some serial correlation over time in the transitory shocks. Such correlation can be tested by calculating the covariance of observed wage outcomes. However, all these papers do not model worker’s job-job selection. Job mobility is arguably the main contribution to transitory shocks (Gottschalk and Moffitt, 1994).
occur quickly. For these reasons, I have assumed that all shocks at match level are permanent ex-ante in worker’s information set.

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

\[
\ln w_{i0} = Z_i' \beta + a_{ij0} + u_{i0} \\
a_{ij0} \sim N(0, \sigma^2_{a0}) \\
E(u_{i0}) = 0, \var(u_{i0}) = \sigma^2_{u0}
\]

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 valued between the worker and the firm. At the beginning of each subsequent period they make 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 search is costless. However, when workers switch jobs, they have to pay a one-time switching cost. Switching cost captures unobserved non-wage factors that drive job mobility choice, which may include, for example, relocation cost and any fringe benefits a worker has to give up if moving to a different job.

I focus on worker’s job-job transitions and ignore 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-job selection process. Hence it does not affect the parameters identified from direct job-job transitions and wage outcomes. Second, from an empirical point of view, job-job transition is the dominating phenomenon for workers in the US labor market compared with transition from employment to unemployment. For working-age male workers, my calculation shows that job-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 young male workers which structural estimation of the model is based upon, 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.

Individuals maximize the expected value of the discounted sum of a time-separable utility function

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

where $\Gamma$ is the discount factor and $E_t$ is the expectations operator conditional on information available in period $t$. $k_{is}$ denotes one-time utility loss in period $s$ if worker switches jobs ($M_{is} = 1$). Assume $k_{is} = k_i + \kappa_{is}$, where $\kappa_{is} \sim N(0, 1)$. $\kappa_{is}$ is a serially uncorrelated shock to switching cost, which introduces a new type of risk into the worker’s decision problem. Individual’s utility function is assumed to be $u(w) = \ln(w)$. As it becomes evident later, the functional form of the utility function is inconsequential to individual’s job-job choice rules. Hours of labor supply is assumed exogenous and inelastic. Wage evolves according to the wage process specified before.\(^{22}\)

Figure 1 depicts the process of job mobility decisions. At the beginning of each period, worker receives an offer from a different firm with probability $\lambda_e$.\(^{23}\) 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 individual and match component of wages, and transitory shock. In the discrete time model, 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 $S(t)$ 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:\n\[ V_{e}^{t}(S(j, t)) = \begin{cases} V_{e}^{t}(S(j, t), M_{t} = 0), & \text{if no offer arrives in period } t \\ \max\{V_{e}^{t}(S(j, t), M_{t} = 0), V_{e}^{t}(S(j', t), M_{t} = 1)\}, & \text{if offer arrives from another firm } j' \text{ in period } t \end{cases} \]

where
\[
V_{e}^{t}(S(j, t), M_{t} = 0) = \ln w_{jt} + \Gamma(1 - \lambda_{e})E_{t}\left[V_{e}^{t+1}(S(j, t+1)|S(j, t))\right] + \Gamma \lambda_{e}E_{t}\left[\max\left\{V_{e}^{t+1}(S(j, t+1), M_{t+1} = 0|S(j, t)), V_{e}^{t+1}(S(j', t+1), M_{t+1} = 1|S(j, t))\right\}\right]
\]
\[
V_{e}^{t}(S(j, t), M_{t} = 1) = V_{e}^{t}(S(j, t), M_{t} = 0) - k_{t}
\]

The set of state variables $S$ includes all variables that provide information on worker’s current and future wages and mobility decisions, including $\{a_{jt}, c, k_{t}\}$. To signify the stochastic evolvement of match-specific wage component in the state variable, set $S$ is indexed by $j$. $V_{e}^{t}(S(j, t), M_{t})$ represents the choice specific value function, for the felicity function depends on $M_{t}$. Similar to most optimal stopping problems, optimal solution is characterized by a threshold reservation value where a worker chooses to move if the offered match is larger than the threshold. In the next two subsections, I discuss the solution under two specifications of the model.

### 3.2.1 Case 1: No Job Switching Cost

If $k_{t} = 0$, then $V_{e}^{t}(S(j, t), M_{t} = 0) = V_{e}^{t}(S(j, t), M_{t} = 1)$, for current period utility can be written as $\ln w_{jt}$ regardless of mobility choice in $t$. Since person component and match component evolve independently and are linearly additive, I decompose $V_{e}^{t}(S(j, t))$ into the sum of $V_{e}^{t}(S(j, t)^{o})$ and $U(t)$ 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.²⁵ The

---

²⁴I omit the person subscript $i$ through the rest of this section.
²⁵The intuition is that worker’s mobility decision is based on $\Delta V = V_{e}^{t}(S(j, t), M_{t} = 0) - V_{e}^{t}(S(j', t), M_{t} = 1)$. $U(t)$ is canceled out in $\Delta V$ and hence state variables other than $S(j, t)^{o}$ do not affect worker’s behavior.
worker’s problem can be re-normalized into

\[
V^e_t (S(j,t)^a) = a_{jt} + \Gamma (1 - \lambda^e) E_t \left[ V^e_{t+1} (S(j, t + 1)^a | S(j, t)^a) \right] + \Gamma \lambda E_t \left[ \max \left\{ V^e_{t+1} (S(j, t + 1)^a, M_{t+1} = 0 | S(j, t)^a), \right. \right.
\[
\left. \left. \left. \left. \left. V^e_{t+1} (S(j', t + 1)^a, M_{t+1} = 1 | S(j, t)^a) \right\} \right. \right. \right. \right. \right.
\]

where the subset of state variables \( S(j,t)^a \) includes \( \{a_{jt}, c\} \). This normalization reduces the number of state variables. Proposition 1 below establishes that, under the assumption of the wage process and zero switching cost, worker is myopic when making job-job transition decisions.

**Proposition 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.** See Appendix A.

Proposition 1 is similar to the reservation wage property in the canonical on-the-job search model, where workers choose to switch jobs if offered wage is higher than current wage (Burdett, 1978). The additional feature of the model–wage uncertainty does not change workers’ job mobility decision. This result is not surprising given our assumption of the wage process. A job offer from another firm and worker’s match with current firm differ only in their levels: worker-firm level wage risk is the same across all firms, and deterministic wage growth is person specific and not firm specific. Thus without switching cost, the continuation value between staying with the current firm and moving to the other firm is the same conditional on \( a_{jt} \). Worker’s decision is only based on a comparison between the contemporaneous utilities from two firms.

### 3.2.2 Case 2: With Job Switching Cost

In light of nonzero switching cost, the reservation match is implicitly defined by:

\[
V^e_t (S(j,t), M_t = 0) = V^e_t (S(j',t), M_t = 1)
\]

The expected value of a job offer must exceed the current value of the job plus the cost of switch. Re-normalizing the agent’s problem as we did for the first case, the set of state variables again can be
reduced to $S(j,t)^a = \{a_{jt}, c, k_t\}$.

Let $h_t(a_{jt}, k_t)$ denote the reservation match in period $t$ conditional on worker’s type $c$. Proposition 2 below establishes that the optimal strategy of job mobility is to move if the value of an offer is larger than worker’s reservation match value.

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

**Proof.** See Appendix A. 

The first and the last properties of the reservation match are direct implications from the monotonicity of the value function. In light of positive switching cost, worker’s reservation match value is always greater than that defined in Proposition 1 which equals to the match paid by the present firm. Positive switching cost is another type of labor market inefficiency besides search friction. 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. Larger cost of switching jobs leads to higher reservation match.

The second property characterizes the dynamic impact of switching cost on job mobility behavior. Let us term the difference between $h_t(a_{jt})$ and $a_{jt}$ as “reservation premium”. Conditional on $k_t$, the convexity of the value function implies that the reservation premium is smaller as $a_{jt}$ grows (the second inequality of the second property). The economic intuition is that worker whose match is low expects more job changes in the future. For these workers, conditional on $k_t$ 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 premium “large”, they avoid paying too much switching cost before reaching a high wage level.\textsuperscript{26}

\textsuperscript{26}Earlier papers analyzing the dynamic effect of switching cost in search models are Hey and McKenna (1979); Van Der Berg (1992). Both 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 on-the-job wage changes. Basic results in Hey and McKenna (1979); Van Der Berg (1992) still hold given our assumed wage process.
3.3 The Role of On-the-job Search: Partial Self-insurance and Identifying True Wage Risk

This section argues how on-the-job search could alter the persistence and magnitude of wage shocks, and how it could arise as a channel of insurance. Suppose a worker’s wage consists of only match-specific component which is subject to permanent shocks. Let us focus on wage risk so shocks to switching cost are ignored in this section. Prior to time $t$, her wage is $a_0$. Left panel of Figure 2 describes an employed worker’s dynamics of match specific wage. At the beginning of period $t$, she suffers a permanent negative match specific shock $\eta$, and her 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, her wage is expected to remain at $a_1$ as long as she stays employed because the shock is permanent in nature.

I consider two scenarios. First, suppose a job offer valued $a^o$ arrives at $t + 1$. Whether the worker takes the offer depends on her reservation match $h(a_1)$ which must be at least as large as $a_1$ (assuming that $k > 0$). If $a^o > h(a_1)$, then she would change job and earn wage $a_2 = a^o$ for it gives her a higher expected wealth. As brought up by Low, Meghir, and Pistaferri (2010), wage increase from $a_1$ to $a_2$ results from an endogenous job mobility rather than wage risk. Moreover, by changing jobs, the worker manages to turn the initial permanent wage shock($\eta$) into effectively partly transitory and partly permanent. Only for some worker who remains at $a_1$ for a long time, the initial shock is correctly identified. The ex-post(observed) persistence of the shocks depends on how quickly a worker could improve her 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 existing estimates from the error component model.\footnote{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.}

Right panel of Figure 2 depicts a second match dynamics in a similar setting. The only difference
is that 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. Observed average wage per period alone mitigates initial wage risk facing workers, as it is combined with worker’s response to latent shocks. Variance of permanent match level shock, $\sigma^2_\eta$, measures wage risk prior to self-insurance.

The welfare costs from removal of an insurance channel are typically measured as the difference between the value functions in the two environments—with and without the insurance channel. This measure, however, overstates the value of insurance because job mobility improves worker’s welfare even in a world without match-specific shocks. If $a_2 > h(a_0)$, then wage gain from $a_0$ to $a_2$ would have taken place even without wage risk. The insurance value of job mobility measures changes in welfare that is affected by shocks. The preferred measure of insurance hinges on how match-level shock affects worker’s job mobility decision and welfare. In our example, if $a_2 < h(a_0)$, worker would not have switched employers if her wage hadn’t decline from $a_0$ to $a_1$. Then wage rise from $a_1$ to $a_2$ measures insurance value of job mobility. If $a_2 > h(a_0)$, worker would have changed employers even if $\eta = 0$. In this case, the insurance value measures wage increase from $a_1$ to $h(a_0)$. In Section 7, I formally define and quantify job mobility as a means of insuring shocks in the labor market.

The preceding discussions demonstrate the economic importance to model the dynamics of match-specific wage $a_j$ and person-specific wage $u$ separately. Although both person- and match-level shocks represent wage risk, they differ in their channel of insurance: permanent shocks at person level cannot be insured through job-job transition, for jobs only differ in the value of match. If $a_j$ is constant within a job tenure, then job mobility never arises as a channel of partial insurance against wage shock and there are no latent wage shocks per se in that context.

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. For every primary and secondary job that respondent holds, SIPP assigns a unique job ID and records job-specific monthly earning. These features, together with a relatively short recall period, make it an attractive data set to study short-term job mobility and wage dynamics.\(^{28}\)

I focus on primary job, which is defined as the job generating the most earnings in a wave. Although SIPP has monthly information on job change 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 earning and wage is derived by deflating the reported monthly earning and wage by monthly US urban CPI. Reported hourly wage rate is used whenever worker is paid by hour. For these workers, 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.\(^{29}\) Job change is identified from a change in job ID between 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 female, 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’s likely that he has already worked at a job for certain periods. In the first wave of SIPP, respondents are asked of 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, worker’s job tenure and wage information is available through the sample period.

As explained in Section 5, I solve the initial condition problem by simulating the model from the first period in work life. This requires one to observe the complete employment histories for any given

\(^{28}\)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 underestimate the extent of job-job transitions by about a fifth.

\(^{29}\)For each month, respondent reports hours of work per week and how many weeks worked. Monthly labor supply is calculated as hours per week \(\times\) (weeks worked/weeks in month) \(\times\) 4.33
worker. As an approximation, I consider job starting between age 20 and 26 as worker’s first job in the life cycle. This leads me to select workers whose calendar age is between 20-26 at the time when their present job started, in addition to the sample restriction described previously. I further restrict the sample to include young workers aged at or below 30 when they first enter the sample. I then keep workers who are included in SIPP for at least eight waves, and construct panel of 1,211 workers whose wages are observed for eight periods.

Summary statistics are provided in Table 1. Table 2 reports the distribution of total number of observed job changes in the sample. 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 switches jobs at least once in the four-year sample period. The extent of job-job transitions decrease monotonically with workers’ labor market experience. There is also some evidence that within-job wage growth slows down with experience. To investigate the influence of observed heterogeneity on the pattern of job-job transitions conditional on experience, I estimate a Tobit model of total number of job changes using all workers. Table 3 shows the results. Conditional on experience, there is strong evidence that workers who own a house make more job-job transitions.

### 4.2 Wage Growth and Job Mobility

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

---

30 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.

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

32 Note that the panel is essentially unbalanced and longer than eight periods since we are using job mobility histories of each worker.
Figure 3A depicts unconditional distribution of within- and between-job real wage growth. Wage growth is calculated as changes in log real wages in every four months\(^{33}\). Two features of the picture are eminent. Firstly, 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. Secondly, both within-job and between-job real wage cuts are very common. Around 45% of job-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 at merely 1.3% per period. There remains, however, a substantial portion of within-job wage growth showing significant drop. 15% of the workers report within-job real wage decline of 11% or more between waves. Wage cut between jobs is 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 keeping up with rising cost of living index in the short run. Figure 3B shows the distribution of between and within nominal wage growth. While the majority of the workers experience nominal wage growth, the amount of workers that had a nominal wage cut remains substantial\(^{34}\).

Column (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 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 up with the expected signs. These two regressions will be run 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. In specific, 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 two-period lagged level of wage and lagged experience as explanatory variables. Columns (4)

\(^{33}\)Throughout this section, wage refers to real wage unless noted otherwise.

\(^{34}\)Workers who are paid by hour experience less frequent nominal wage cut within jobs (yet remains sizable: about 20% within job wage changes are wage cuts), than workers who are paid by salary.
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 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 three sources: from reported wages for those who are hourly paid, and from reported earnings and hours for salary-paid workers. Assuming that these reporting errors are classical (i.i.d.), the extent of wage cut documented previously may be an overstatement of the true wage changes. Gottschalk (2005) uses earlier waves of SIPP and estimates that as many as three quarters of observed decline in within-job nominal wages reflect 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 fairly large (Section 6), suggesting that measurement error indeed has a large impact on cross-sectional variations 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 errors, there should be zero correlation between worker’s job mobility choices and wage movement on the job. However, probit regressions in Table 5 demonstrate strong correlation between observed within-job wage change 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 errors, studies typically find little effect of nonclassical measurement error on earning mobility and 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 for future research.

I view the empirical results documented above as strong evidence for job mobility as a responsive behavior to wage risk. It is useful to reiterate that these empirical relations do not carry any causal interpretation. I now turn to structural estimation of the dynamic model of job mobility, where job-selection mechanism is implied by utility-maximizing rational workers.

5 Identification and Estimation Strategy

The estimation consists of two steps. In the first step, I estimate the earning regression and produce a consistent estimator of $\beta$. Recognizing the analytical complexity of evaluating conditional expectations in theoretical moments, the second step uses Method of Simulated Moments (MSM). The MSM numerically simulates the expectation of wage and job mobility, and then minimizes the distance between the simulated moments and empirical moments derived from data\textsuperscript{35}.

5.1 The Empirical Model

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

$$k_{it} = X_i'\Upsilon + \kappa_{it}$$

(12)

Vector $X$ includes a constant, marital status, whether own a house, two education dummies(high school, at least some college) and dummy on race. As defined previously, $\kappa_{it}$ is the i.i.d random shock to switching cost with mean zero and variance normalized to one. It represents non-wage risk against which job mobility could potentially insure. $\Upsilon = \{\gamma_{constant}, \gamma_{married}, \gamma_{house}, \gamma_{hs}, \gamma_{college}, \gamma_{white}\}$ is the

\textsuperscript{35}The primary reason for choosing MSM over a maximum likelihood estimator is that, to construct the likelihood function, one needs to integrate out the person component of wage $u$ in every period. As the illustration below goes, the first- and second-moments of person-specific wages carry a tractable analytical form that is easy to evaluate.
vector of coefficients to be estimated\textsuperscript{36}.

Next, let us lay out the empirical model for two specifications of the search model. Recall that the wage equation for an individual $i$ employed by firm $j$ in his labor market age $t$ is:

$$
\ln \bar{w}_{ijt} = \ln w_{ijt} + v_{it}
$$

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

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

$$
a_{ijt} = \begin{cases} 
    a_{ijt-1} + c_i + \eta_{ijt} \equiv a_{ijt}^l, & \text{if } M_t = 0 \\
    a_{ij't}, & \text{if } M_t = 1
\end{cases}
$$

$$
u_{it} = u_{it-1} + \delta_i + \zeta_{it}
$$

where as previously defined, $a_{ij't}$ is the value of random offer, $a_{ij't}^l$ denotes latent match-level wage in period $t$ prior to job mobility decision, and $M_t$ is the observed mobility indicator, which equals to one when there is an acceptable offer. Without switching cost, job mobility decision can be formulated as a bivariate selection model:

$$
J_{it}^* = a_{ij't}^o - a_{ijt}^l 
$$

$$
J_{it} = \begin{cases} 
    1 \text{ if } J_{it}^* > 0 \\
    0 \text{ elsewhere}
\end{cases}
$$

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

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

where $J_{it}^*$ is the offer-acceptance rule (choice); $O_{it}$ defines the outcome of a trivial offer-arrival process.

\textsuperscript{36}Each of $X_i$ in $X$ is measured in the first observation period and is assumed time-invariant starting from the beginning of life. One could solve the dynamic programming problem for each worker given his observed $X_{it}$ and expected $X_{it}$ beyond the observation period. This would, however, dramatically increase the computation burden, as opposed to solving the model for each type of worker holding $X_i$ fixed over life. A more difficult question is to allow marriage and house ownership to be endogenously determined. I leave it for future research.
(chance) which is common across workers and independent to \( J_{it} \). \( J_{it}^* \) 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. Without switching cost, the job selection rule is simply based on the difference between offered match and current match value (Proposition 1). Evaluating the selection equation does not require one to solve the value function, and it is equivalent to solving a static model for every period.

With nonzero switching cost, the only difference is the offer acceptance rule:

\[
J_{it}^* = (V_t^c(S(j^t, t)^a, M_t = 0) - k_{it}) - V_t^c(S(j, t)^a, M_t = 0)
\]  

(17)

The state variable vector contains \( \{X_i, a_{ijt}, c_i, \kappa_{it}\} \). With switching cost, we need to solve the value function for every possible combination of the state variables. The computational difficulty arises because \( a_{ijt} \) and \( c_i \) are continuously distributed and serially correlated unobserved state variables. State variable \( c_i \) represents unobserved heterogeneity to the econometrician (although each worker knows his \( c_i \)). For estimation purpose we need to solve the value function for different value of \( c_i \). Furthermore, to solve the finite horizon dynamic programming problem at \( t \), we need to know the value function at \( t + 1 \) for every possible combination of state variables \( a_{ijt+1} \) that may arise conditional on \( a_{ijt} \). As \( t \) increases, the problem quickly becomes intractable\(^{37}\). Appendix B describes the solution method in detail. The method uses Monte Carlo integration and an interpolation method to approximate the value function.

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. As shown later, non-wage factors are very important determinants in job mobility decisions.

### 5.2 Identification

Our preferred model is parsimoniously characterized in terms of the parameter vector\(^{38}\)

\[
\Theta = (\beta, \lambda^e, \sigma_\delta^2, \sigma_c^2, \sigma_{u_0}^2, \sigma_{\eta_0}^2, \sigma_\gamma^2, \sigma_v^2, \sigma_\mu^2, \mu_c, \mu_\delta, \Upsilon)
\]  

\(^{37}\)This is known as the “curse of dimensionality” problem. See Aguirregabiria and Mira (2009) and the reference therein.

\(^{38}\)The discount factor \( \Gamma \) is held fixed at 0.97.
Coefficients ($\beta$) on observable vector $Z_i$ is identified by the initial wage assumption. Since wage residual observed in the first period of work life is assumed exogenous (no selection takes place prior to the beginning of work life), OLS regression of observed wages in the first period work life produces consistent estimates of $\beta$.

The preceding discussion demonstrates that job mobility choice implements 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-selec- tion rule. In the full model with switching cost, the job-selection rule can be thought of as a reduced-form equation on $X_t$, which is analogous to the choice equation in the classic Heckman selection models. Therefore, the switching cost parameters ($\Upsilon$) 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.\textsuperscript{39}

I partition the rest of the parameters into two sub-vectors: $\Theta^m = \{\lambda^e, \sigma_{a0}^2, \sigma_\eta^2, \sigma_\zeta^2, \mu_\zeta, \mu_\delta\}$ and $\Theta^n = \{\sigma_{u0}^2, \sigma_\zeta^2, \sigma_\nu^2\}$. $\Theta^m$ contains parameters governing the evolution of match-level wage and the search process. $\Theta^n$ 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. From Appendix D, it’s clear to see that covariance of wage and mobility, autocovariance of mobility, mean of mobility and wage only depend on $\Theta^m$. $\Theta^m$ is constant over time. With observations from sufficient periods, these moments overidentify all the parameters in $\Theta^m$.\textsuperscript{40} 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 assumption on the shocks and random growth factor in the match-level process is essential in evaluating the mass at the bottom of the distribution.

Given $\Theta^m$, $\Theta^n$ can be identified using moments based on variance and autocovariance of wage alone, like most papers in the wage dynamics literature. No distributional assumptions are required to identify parameters governing the person-level wages. $\sigma_\delta^2$ and $\sigma_\zeta^2$ impose that the autocovariance

\textsuperscript{39}As a sensitivity test, I also estimate the model assuming zero switching cost, whose identification does not depend on the exclusion restriction. The estimated wage parameters are similar (see Section 6).

\textsuperscript{40}Note that, in estimation, the parameters are estimated jointly imposing all moments simultaneously.
of person-level wages grow nonlinearly and linearly respectively with time locations (see formula in Appendix D). Hence they are identified by fitting in 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 $\sigma^2_u$ is identified from the variance of wages at the beginning of work life.

Lastly, I provide some intuition as to where identification comes from, for parameters in $\Theta^m$. As is evident from equation (11) in Appendix D, the mean of return to experience ($\mu_\delta$) is identified by fitting a linear trend in the observed wage residuals over life. Within-job wage growth over life identifies return to tenure and the combination of match-level uncertainty ($\sigma^2_\eta$) and heterogeneity in return to tenure ($\sigma^2_c$)\footnote{Within-job wage growth alone and first moment of wage in general do not separately identify $\sigma^2_\eta$ and $\sigma^2_c$ because they are linearly additive in $\sigma^2_{a_0}$. 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.}. 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, empirical evidence suggests that within-job wage growth slows down with experience (Table 2). By modeling the dynamics of 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 $\sigma^2_\eta$ (or $\sigma^2_c$) changes the slope and curvature of within-job wage growth profile over time and different $\mu_c$ only changes its intercept. Therefore, the slope and curvature of the fitted profile of within-job wage growth identifies the combination of $\sigma^2_\eta$ and $\sigma^2_c$, and, given $\mu_\delta$, its intercept identifies $\mu_c$.

To separately identify $\sigma^2_c$ and $\sigma^2_\eta$, 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 $\sigma^2_\eta$ and $\sigma^2_c$ while holding the rest of the parameters fixed. While both greater $\sigma^2_\eta$ and $\sigma^2_c$ lead to an increase in the covariance of mobility across different lags, the relative effect from $\sigma^2_\eta$ is quickly deteriorating with the length of the lag. Hence $\sigma^2_c$ can be separately identified from $\sigma^2_\eta$ using autocovariance of mobility at sufficiently long lags.

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

5.3 The Initial Condition Problem

In the model, a worker begins his first job with a random draw from the offer distribution. Therefore, wages in the first period of work life are exogenous. Observed wages in subsequent periods are as a result of selection and, therefore, are endogenous.

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 observed wages are endogenous and we have an initial condition problem (Heckman, 1981). Recall from Section 4.1 that 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 \( \tau \) 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 \( \tau \) together map into a unique work life period \( t: p + \tau = t \).

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 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 statistic (e.g. mean wages) in period \( \tau + p \) conditional on \( M_0 = M_1 = ... = M_\tau = 0 \) is a consistent estimator for the observed statistic in observation period \( p \).\(^{42}\)

5.4 MSM Estimation

Due to computational complexity, it is difficult to accurately estimate all elements of \( \Theta \) in one step. I employ a two-stage estimation procedure. In the first stage of the estimation, I select workers whose initial wage on their first job is observed\(^{43}\). As discussed previously, regressing it on observed personal

\(^{42}\)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.

\(^{43}\)672 workers satisfy this criteria.
characteristic $Z$ yields a consistent estimator of $\beta$ - the coefficient vector on $Z$. Vector $Z$ includes a constant, education dummies, race, a polynomial in calendar age and interactions between age and education dummies. *Age* here captures the effect of potential labor market experience on worker’s initial wage. Consistent with 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 variations to unobserved heterogeneity in match- and person-level wage component.

Given a consistent estimate of $\beta$, I obtain predicted log wage residual for each worker in each observation period in the sample. 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 ($\mu_c$), return to experience ($\mu_\delta$) and worker’s selection into better jobs. For the remaining analysis I work with log wage residuals denoted by $r_{ip}$.

Remaining parameter vector, denoted by $\theta$, is estimated by 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 $s_i(M_{p1}, r_{p1}, M_{p2}, r_{p2})$ for any two observation periods $p_1$ and $p_2$ such that $1 \leq p_1 \leq p_2 \leq P$:

$$s_i(M_{p1}, r_{p1}, M_{p2}, r_{p2}) = [M_{ip2}, r_{ip2}, (r_{ip2} - \overline{r}_{p2})(M_{ip1} - \overline{M}_{p1}), (M_{ip2} - \overline{M}_{p2})(r_{ip1} - \overline{r}_{p1})]$$

$$= (r_{ip2} - \overline{r}_{p2})(r_{ip1} - \overline{r}_{p1}), (M_{ip2} - \overline{M}_{p2})(M_{ip1} - \overline{M}_{p1}))'$$

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

The empirical model to be estimated is:

$$s_i(M_{p1}, r_{p1}, M_{p2}, r_{p2}) = f(\theta; X_i, \tau_i, p_1, p_2) + \varepsilon_i$$

Function $f$ is the expected value of $s_i(M_{p1}, r_{p1}, M_{p2}, r_{p2})$ as implied from the model. It is important to note that the predicted moment $f$ depends on $\tau_i$ (elapsed duration of the first job at $p = 1$), because of the left-censoring problem in the sample. $\tau_i$, $p_1$, and $p_2$ map into two unique life periods $t_1$ and $t_2$. Conditional on observed state variables, the theoretical moments is a function of life period $t_1$ and $t_2$. 

28
conditioning on that $M_0 = M_1 = \ldots = M_{\tau_i} = 0$.

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:

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

(21)

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. $\{\hat{\nu}_s\}_{s=1}^{S}$ 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 $\eta$ 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. Person-component variables are additive to match-component and they 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 $\text{var}(r_{it})$ consisting of a linear combination of $\text{var}(u_{it})$ and $\text{var}(a_{ijt})$. As person-level wage $u_{it}$ is independent of job mobility decision, the former can be expressed as function of model parameters:

$$\text{var}(u_{it}) = \sigma^2_{u_0} + t_2^2 \sigma^2_{\eta} + t_2 \sigma^2_{\eta}.$$ 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, LHS of equation (20) becomes the mean, variance and covariance of mobility and wage residuals between any two observation periods in the sample:

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

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

---

44 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 gradient-based optimization method. I smooth $M_{it}$ by $\Phi(J^*_{it} h)$, where $\Phi$ is the standard normal cumulative distribution function and $h$ is a smoothing parameter. When $h \to 0$, the approximation converges to the frequency simulator.
The vector of simulated moments for any given two observation periods is:

$$g(\theta; p_1, p_2) = s(M_{p_1}, r_{p_1}, M_{p_2}, r_{p_2}) - \frac{1}{N} \sum_{i=1}^{N} \hat{f}(\theta; X_i, \tau_i, p_1, p_2)$$  \hspace{1cm} (22)$$

where $N$ is the number of workers in the panel, $s$ 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 $X, Z, \tau$ as in the observed sample.

Let $g(\theta)$ be a vector consisting $g(\theta; p_1, p_2)$ at all possible combinations of $p_1$ and $p_2$. The size of vector $g(\theta)$ 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$^{45}$. The goal of the second stage estimation is to find $\theta$ which minimizes:

$$g(\theta)'Wg(\theta)$$  \hspace{1cm} (23)$$

where $W$ is an M by M weighting matrix. The weighting matrix is chosen as a diagonal matrix such that $W = A^{-1}$. $A$ is a diagonal matrix whose elements on the main diagonal are given by

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

Therefore, the main diagonal of $W$ contains the inverse of the variance of corresponding elements in $s$. 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$^{46}$. 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)).

$^{45}$Beyond eight sample periods, the sample size is not large enough to yield reliable estimates of empirical moments.

$^{46}$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.
6 Estimation Results

Estimated parameters are presented in Table 6. The first column contains estimates for our preferred model with switching cost. Wage risk at worker-firm match level is the dominating risk facing employed workers: the variance of match-level shock \((\sigma_\eta^2)\) is \(0.829\) \text{100}, which is more than two orders larger than the variance of person-level wage shock \((\sigma_\zeta^2 = 0.005\) \text{100}). It implies that on-the-job search could potentially insure a huge fraction of permanent wage risk. Variance of 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 \((\mu_c)\) is about -0.5% per period (four months) and the estimated return to experience \((\mu_\delta)\) 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 shock, both of which are determinants of job mobility choice and are absent in reduced-form estimation of return to tenure\(^{47}\).

The negative \(\mu_c\) 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-job transitions in the data. Our story is similar here: if there is no depreciation in job-specific skills (i.e. \(\mu_c\) is positive), it implies a quick decline in the probability of moving over time which is inconsistent to the high rate of job-job transitions we observe even for workers. Note that even though mean return to tenure is negative, selection effect over the level and the change in match are able to generate sufficient match improvement over time.

The estimated heterogeneity of the individual wage component at the start of life is much larger than initial match heterogeneity, suggesting that the individual wage component is essential to match both the extent of job-job transitions and the associated dispersion in wages\(^{48}\). Examining the switching cost parameters, we find that estimated switching cost is moderately large. The magnitude of estimated \(\gamma_{\text{constant}}\) is more than one-third of the mean log wage in the first period of life. For this sample of young male workers, switching cost is smaller for workers who have a college degree, own a house, and

\(^{47}\)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.

\(^{48}\)Bils, Chang, Kim, and Hall (2009) and Hornstein, Krussell, and Violante (2007), show that match heterogeneity alone is insufficient to produce both realistic wage dispersion and unemployment fluctuations at the same time.
for married and white workers. 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 their mortgage payments.

What happens if shocks to 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 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 huge increase in the estimated variance of permanent shock (from $0.005^{100}$ to $0.176^{100}$). A large proportion of wage fluctuations that is in fact specific to a worker-firm match have been identified as permanent shocks that will persist across all jobs. In Section 7, I discuss the implications of this finding to the true wage risk facing workers. In column (4), I estimate a canonical wage process by neglecting match-specific wage component and worker’s selection between jobs. This has been a standard in estimating wage uncertainty in 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 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 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-job transitions and wage growth from data.

---

49 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 ($O$).

50 Since wage is assumed exogenous (no selection between jobs), the first stage regression is run using observed wages over all periods on personal characteristics including labor market experience. Transitory shocks are i.i.d although this assumption can be relaxed. Following the literature, the weighting matrix is set as an identity matrix.
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 $\tau$. The final simulated sample contains 8 observation periods, whose joint distribution of $X, Z$ and $\tau$ 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.

7 Implications of the Model

7.1 Decomposing Wage Growth and Wage 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 the contribution from match component and individual component of wages overtime. The left panel of Figure 6 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 selection, since the mean return to tenure is negative. Good shocks are preserved and bad shocks could be recovered through job changes. As the extent of job-job transitions decreases with experience, the growth of $E(a_t)$ gradually slows down, which explains the concavity of the experience profile of wages.

The right panel of Figure 6 decomposes the experience profile of wage inequality. At the beginning of life, almost all wage inequality is from variations in individual heterogeneity. As labor market experience grows, the contribution from worker-firm match quickly rises because of permanent match shocks, job-job transitions and heterogeneous return to tenure. After 10 years since the beginning of life, variation at match level is about the same magnitude as wage variation at person level. Notice that
measurement error plays a significant role in explaining cross-sectional wage inequality throughout life.

7.2 Wage Risk Prior to Job Mobility

True wage risk corresponds to the wage risk prior to making job mobility decision in each period which is in worker’s information set but is unobserved to the econometrician. In our model, the true wage risk is the sum of the variance of person- and match-level shock ($\sigma^2_\zeta + \sigma^2_\eta = 0.829 + 0.005$), which is more than twice as much as the variance of permanent shock identified from a canonical wage process where firm-specific wage is neglected (0.321). When switching cost is ignored, the estimated true permanent uncertainty (0.624+0.003) 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 worker’s reservation match and there must be larger variations in the match-level shock in order to explain the observed mobility patterns from data.

Another way to argue that wage risk is underestimated from realized wage change is to compare the variance of realized match-specific wage change with the variance of match-level shock. Figure 7 plots the variance of changes in match qualities 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 dynamic 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. Variation of realized match changes (dotted line) underestimates the true wage variation (solid line) which would have taken place without job mobility. Higher switching cost (dashed line) reduces 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 Insurance Value of On-the-job Search

In this section, I formally define and quantify the insurance value of on-the-job search. On-the-job search potentially insures two types of shock: match-level wage shock and unobserved non-wage shock to job switching cost. Two measures of insurance are introduced. The first and the preferred measure is based on the difference between the discounted expected life-time utility at the beginning of life with
free job mobility and a properly defined counterfactual value function. This measure of insurance compares how welfare is affected by shocks in the presence of on-the-job search. Define the value of insurance $\Delta V_0$ at the beginning of life as

$$\Delta V_0 = V_0^c(S(j, 0)) - V_0^c(S(j, 0))$$  \hspace{1cm} (24)$$

where $V_0^c(S(j, 0))$ is the value of a job in a hypothetical environment where match-level shocks and shocks to switching cost are shut down when a worker makes job mobility choice. In specific, the counterfactual value function is evaluated with respect to a modified conditional density function of match $f'$ in every period. Conditioning on worker’s type and history of match draws up to $t - 1$, let us define the density function of match $f'(a_{ijt}|a_{ijt-1}, \eta_{it})$ at $t$ as

$$f'(a_{ijt}|a_{ijt-1}, \eta_{it}) = g(a_{ijt}|M_{it}, a_{ijt-1}, \eta_{it}) h(M_{it}|a_{ijt-1}, \eta_{it} = 0, \kappa_{it} = 0)$$  \hspace{1cm} (25)$$

where $\eta_t$ is absent but only for the $h$ density, not the $g$ density.

Table 7 shows the calculated the insurance value for a group of workers differing by their mean switching cost. The insurance value $\Delta V_0$, if represented as a percentage of the value a job $V_0^c$, is nearly 25% for all types of workers. Looking at magnitude of $\Delta V_0$ for different types of workers (column 2), we find that workers whose switching cost is low (e.g. married, do not own a house, college graduate, white workers) have a larger insurance value 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). Intuitively, conditional on the distribution of match draws, high-switching-cost workers choose a high reservation wage which makes both job-job transition an unlikely event and outside job offers less attractive.

For the second measure of insurance, define $\xi_t(S(j, t), \eta_{jt})$ as the compensating utility to make a worker indifferent between having and not having a negative match-level shock:

$$V_t^c(a_{jt} - \eta_{jt}, c, k_t) + \xi_t = V_t^c(a_{jt}, c, k_t)$$  \hspace{1cm} (26)$$

$\xi_t(S(j, t), \eta_{jt})$ evaluates how much utility we would need to transfer to a worker conditional on a

---

51If counterfactual value function is defined where job mobility is completely shut down, then it overestimates the value of insurance because job mobility improves welfare even in a world without match-specific shocks.
particular $\eta_{jt}$. Now consider another environment where job mobility is disallowed. Analogously, to make the worker indifferent between getting and not getting the same $\eta_{jt}$, the compensating utility $\tilde{\xi}_t(S(j, t), \eta_{jt})$ is

$$V_t^e(a_{jt} - \eta_{jt}, c, k_t; \lambda^e = 0) + \tilde{\xi}_t = V_t^e(a_{jt}, c, k_t; \lambda^e = 0)$$ (27)

The insurance value of job mobility (conditional on $\eta_{jt}$) can be then calculated as

$$\tilde{\omega}(S(j, t), \eta_{jt}) = \frac{\tilde{\xi}_t(S(j, t), \eta_{jt})}{\tilde{\xi}_t(S(j, t), \eta_{jt})}$$ (28)

If $\tilde{\omega} = 0$, the worker is fully insured; if $\tilde{\omega} \geq 1$, the worker is not insured. In general, we expect job mobility to partially insure a worker against a match-level shock (i.e. $0 < \tilde{\omega} < 1$). The smaller $\tilde{\omega}$ is, the better insured a worker is.

Assume that $\eta_{jt}$ is a one standard deviation negative shock. Given the estimated parameters, I calculate $\tilde{\omega}$ and plot it in Figure 8 for three types of workers differing in their switching costs at the beginning of work life. Similar to the implication from the first measure of insurance, workers with high switching cost are not as well insured as those with low cost of switch. Low-match workers are better insured than high-match workers. The degree of partial insurance is fairly high. For a median worker at the start of his career, $\tilde{\omega} = 0.18$, which indicates that 82% of the potential welfare loss caused by a one-standard deviation match shock can be recovered through changing employers.

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 worker-firm match level and shocks at individual level which persists across jobs. The key identifying restriction is that match-level shocks affect worker’s job mobility decisions, while 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, for two reasons. One is that it enables econometricians to identify the true wage risk prior to job mobility, by using information on both wages and job mobility choices. The other reason is that job mobility could be an important channel for insuring against match-level wage shocks. Results from
this paper suggest that true wage risk doubles the wage risk when estimated using wage information alone. Job-job transition is a very valuable channel of insurance which could insure as much as 82% of the potential welfare loss from a one standard deviation match-level shock for a median worker. The extent of latent wage risk and the value of insurance depend critically on worker’s reservation wage. In the search model, workers are heterogeneous in terms of match quality, switching cost and return to tenure. All of these generate heterogeneity in the reservation wage and in turn produce heterogeneity in the value of insurance and the extent of latent wage risk.

The policy implication from this paper is that, during an economic recession, the welfare of employed workers could also be greatly reduced. Most of the research projects so far focus on the welfare for displaced workers. Yet for employed workers, recessions could lead to a rise in job switching cost or a decline in the offer arrival probability. Under both cases, the insurance value of job mobility decreases. For instance, if job arrival probability drops by a half for a median worker at the beginning of his career, only 73% of the potential welfare loss caused by a one-standard deviation match shock can be recovered through changing employers - a 12% drop. Under our model, in recessions, match-level shocks would appear more persistent and wage risk identified from wage data would be larger.\textsuperscript{52}

This paper can be extended in a number of directions. 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 the job quality is allowed (Jovanovic, 1979), the variation of surprises to the match quality is decreasing with tenure as workers learn more about their match-specific productivity. It is interesting to explore the implication of the model where job is an “experience good”. Also, 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. Additionally, this paper limits job mobility to be the only channel which could insure against match-level risk. Occupation choice, labor supply and savings are other prominent channels which could potentially interact with job mobility choice and insure against wage risk at match level. An important avenue for future research is to analyze the relation between job mobility and other insurance channels, and to quantify their relative importance for insuring against different types of shocks.

\textsuperscript{52}This coincides with empirical evidence from the literature finding that idiosyncratic risk is countercyclical (Storesletten, Telmer, and Yaron, 2004).
APPENDIX

A Proofs of Propositions 1 and 2

In this section I provide formal proofs to Propositions 1 and 2 in the paper. For notational convenience, I make two simplifying assumptions. First, assume switching cost is a constant $k_i$ for every worker. In the paper, shocks to $k_i$ is i.i.d and hence does not affect the value function except for the current period utility. Second, I omit state variables $c_i$ and $k_i$ in the proof, leaving $a_{ijt}$ as the only state variable in worker’s problem. The value function defined here is therefore 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.

$$V_t^e(a_{ij}) = a_{ij}$$

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

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

(1)$$

We know that

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

By the assumptions on the match process, it’s easy to see 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 $E_t[V_{t+1}^e(a_{ijt+1})]$ 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’s easy to see 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 $E_t[V_{t+1}^e(a_{ij't+1}) - k_i]$ is increasing in $a_{ijt}$ and hence $V_t^e(a_{ijt})$ is increasing in its argument.

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

Proof. The reservation match value satisfies:

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

(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\(^{53}\),

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

Lemma 3. \( V_t^e(a) \) is monotonically increasing in \( a \), for all \( t=1\ldots T-1 \).

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

\[
\frac{\partial V_t^e(a_{ijt})}{\partial a_{ijt}} = 1 + \Gamma(1 - \lambda^e)E_t[V_{t+1}^e(a_{ijt+1})']
\]

\[
+ \Gamma \lambda^e E_t \left[ V_{t+1}^e(a_{ijt+1})' \Phi(h_{t+1}(a_{ijt+1})) + V_{t+1}^e(a_{ijt+1}) \phi(h_{t+1}(a_{ijt+1}, k_i)) \frac{\partial h_{t+1}(a_{ijt+1}, k_i)}{\partial a_{ijt+1}} \right]
\]

\[
- \Gamma \lambda^e E_t \left[ (V_{t+1}^e(h_{t+1}(a_{ijt+1})) - k_i) \phi(h_{t+1}(a_{ijt+1}, k_i)) \frac{\partial h_{t+1}(a_{ijt+1}, k_i)}{\partial a_{ijt+1}} \right]
\]

\[
= 1 + \Gamma(1 - \lambda^e)E_t[V_{t+1}^e(a_{ijt+1})'] + \Gamma \lambda^e E_t \left[ V_{t+1}^e(a_{ijt+1})' \Phi(h_{t+1}(a_{ijt+1})) \right]
\]

(3)

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

\[
\frac{\partial^2 V_t^e(a_{ijt})}{\partial a_{ijt}^2} = \Gamma(1 - \lambda^e)E_t \left( \frac{\partial^2 V_{t+1}^e(a_{ijt+1})}{\partial a_{ijt+1}^2} \right)
\]

\[
+ \Gamma \lambda^e E_t \left[ \frac{\partial^2 V_{t+1}^e(a_{ijt+1})}{\partial a_{ijt+1}^2} \Phi(h_{t+1}(a_{ijt+1}) + V_{t+1}^e(a_{ijt+1})' \phi(h_{t+1}(a_{ijt+1}, k_i)) \frac{\partial h_{t+1}(a_{ijt+1}, k_i)}{\partial a_{ijt+1}} \right]
\]

(4)

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,

\[
V_t^e(a_{ijT-1}) = a_{ijT-1} + \Gamma(1 - \lambda^e)E_{T-1} [a_{ijT}] + \Gamma \lambda^e E_{T-1} \left[ \max a_{ijT}, a_{ijT} - k_i \right]
\]

\(^{53}\)\( V'(a) \) denotes partial derivative of the value function w.r.t \( a \).
It is straightforward to show that \( \frac{\partial^2 V^e_t(a_{ijt})}{\partial a^2_{ijt}} = \Gamma \lambda^e E_{T-1}(\phi(a_{ijT} + k_i)) > 0. \)

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

**Proof.** The reservation match value is defined by:

\[
V^e_t(a_{ijt}) = V^e_t(h_t(a_{ijt}, k_i)) - k_i
\]

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

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

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

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 \( \frac{\partial V^e_{T-1}(a_{ijT-1})}{\partial k_i} \) 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 \( k \)

\[
\frac{\partial V^e_t(a_{ijt})}{\partial k_i} = 1 + \Gamma(1 - \lambda^e)E_t\left(\frac{\partial V^e_{T-1}(a_{ijt+1})}{\partial k_i}\right) + \Gamma \lambda^e E_t \left[ \frac{\partial V^e_{T+1}(a_{ijt+1})}{\partial k_i} \Phi(h_{t+1}(a_{ijt+1}, k_i)) \right] \\
+ \Gamma \lambda^e E_t \int_{h_{t+1}(a_{ijt+1}, k_i)}^{\infty} \frac{\partial V^e_{T+1}(a_0)}{\partial k_i} d\Phi(a_0) - \Gamma \lambda^e E_t [1 - \Phi(h_{t+1}(a_{ijt+1}, k_i))]
\]

(6)
By the induction assumption, it is obvious that \( \frac{\partial V^e_{t+1}(a_{ijt})}{\partial k_i} < 0 \). Then, since we know\(^{54}\)

\[
\int_{h_{t+1}(a_{ijt+1}, k_i)}^{\infty} \frac{\partial V^e_{t+1}(a_0)}{\partial k_i} d\Phi(a_0) > \frac{\partial V^e_{t+1}(h_{t+1}(a_{ijt+1}, k_i))}{\partial k_i} (1 - \Phi(h_{t+1}(a_{ijt+1}, k_i)))
\]

Equation (6) then simplifies to

\[
\frac{\partial V^e_{t}(a_{ijt})}{\partial k_i} > \Gamma E_t \left( \frac{\partial V^e_{t+1}(a_{ijt+1})}{\partial k_i} \right) > -1
\]

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

\[
\frac{\partial V^e_{t}(a_{ijt})}{\partial k_i} + 1 > 0
\]

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)\(^{55}\).

The value function does not have an analytical solution because of the assumption on the evolution of state variables. Recall that the state variables contain \( \{c_i, a_{ijt}, X_i, \kappa_{ijt}\} \). The continuation value (for a person of \( c_i \) and \( X_i \) employed by firm \( j \)) can be approximated by Monte Carlo simulations. For

\(^{54}\)This builds on that \( \frac{\partial V^e_{t}(a_{ijt})}{\partial k_i} \) is increasing in \( a \), which can be proved by induction. It is easy to see that this holds for period \( T - 1 \).

\(^{55}\)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 \( T_0 \).
example, the $E_{\text{max}}$ function in equation (1) of the Appendix is given by

$$
\frac{1}{D} \sum_{d=1}^{D} \max(V_{t+1}^e(a_{ijt+1}^d), V_{t+1}^e(a_{ijt}^d) - X_i' \Upsilon + \kappa_{it}^d)
$$

(7)

where $a_{ijt+1}^d, a_{ijt}^d,$ and $\kappa_{it}^d$ is the $d_{th}$ draw from the distribution of $a_{ijt+1}$ given $a_{ijt}$ (equivalent to a draw from the distribution of $\eta_{it}$), $a_{ijt}^d$ and $\kappa_{it}$ 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 $x$ so that $E(x)$ can be approximated by $\frac{1}{n} \sum_{d=1}^{n} x_d$. 56

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 $a_{ijt}^n$. 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 $a_{ijt+1}^{n-1}$ and $a_{ijt+1}^n$. 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 shock57. Both

56 $x_i$ 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 $E(x|A_i^{-1} \leq x \leq A_i)$ $= \int_{A_i^{-1}}^{A_i} x dF(x)$.  

57 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 $\tau_i$. 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 $M_0 = M_1 = \ldots = M_{x_0} = 0$. For this particular worker, the simulated moment $\hat{f}$ is then

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

where $S_{x_0}$ is the number of simulated job histories that are in set $A$.\(^{58}\)

**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 $\Upsilon$ 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 $\sigma_{a^o}^2$. Let $a^t_i$ denote the latent match in period $t$ before job mobility decision is made. Consider the unconditional first moment of mobility and wages:

$$E(M_t) = E(E(M_t|a^t_i)) = \lambda c E_{a^t_i} \left[ \Phi(-\frac{a^t_i}{\sigma_{a^t}}) \right]$$

$$E(r_t) = E(a_t) + E(u_t) + E(v_t)$$

$$= E(a_t|M_t = 1)E(M_t) + E(a_t|M_t = 0)(1 - E(M_t)) + \mu_\delta \times t$$

$$= \sigma_{a^o} E_{a^t_i} \left[ \lambda(a^t_i) \right] E(M_t) + \left[ E(a^t_i)(1 - \lambda^c) + \lambda^c E_{a^o} \left[ E(a^t_i|a^t_i > a^o) \right] \right] (1 - E(M_t)) + \mu_\delta \times t \quad (10)$$

where $\lambda$ is the inverse Mills ratio: $\lambda(a^t_i) = \frac{\phi(a^t_i/\sigma_{a^t})}{1 - \Phi(a^t_i/\sigma_{a^t})}$. $\Phi$ and $\phi$ are standard normal c.d.f and density function respectively. $a^t_i$ is a function of complete history of wage draws: $c, a^o, \{\eta_p\}_{p=1}^t, \{a^o_{j,p}\}_{p=1}^{t-1}$. The

\(^{58}\)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 \( a_t^1 \) is a function of \( \Theta^m = \{ \lambda, \sigma_{a_0}^2, \sigma_\eta^2, \sigma_c^2, \mu_c, \mu_\delta \} \).

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

\[
\text{cov}(r_k, M_t) = E(r_k | M_t = 1)E(M_t) - E(r_k)E(M_t) \\
\text{cov}(M_k, M_t) = E(M_t | M_k = 1)E(M_k) - E(M_k)E(M_t)
\]

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:

\[
E(r_k | M_t = 1) = E(a_k | M_t = 1) + E(u_k | M_t = 1) + E(v_k | M_t = 1) \\
= E(a_k | M_t = 1) + \mu_\delta \times k
\]

\[
E(M_t | M_k = 1) = P(M_t = 1 | M_k = 1) = \frac{P(M_t = 1, M_k = 1)}{E(M_k)}
\]

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 \( \Theta^m \). 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 \( a_t^1 \) is normally distributed with mean \( \mu_c \) and variance of \( \sigma_{a_1}^2 = \sigma_{a_0}^2 + \sigma_\eta^2 + \sigma_c^2 \). The joint distribution of \( a_t^1 \) and \( a_0 \) follows a bivariate normal distribution with means \( \mu_c \) and zero, variances \( \sigma_{a_1}^2 \) and \( \sigma_{a_0}^2 \), and correlation coefficient \( \rho = \frac{\sigma_{a_1} \sigma_{a_0}}{\sigma_{a_1}^2} \).

Then,

\[
E(r_0 | M_1 = 1) = \rho \sigma_{a_0} E_{a^0} [\lambda(a^0)]
\]

where \( \lambda(a^0) = -\frac{\phi((a^0 - \mu_c)/\sigma_{a_1})}{\Phi((a^0 - \mu_c)/\sigma_{a_1})} \).

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

\[
\text{cov}(r_{it}, r_{ik}) - \text{cov}(a_{it}, a_{ik}) = \sigma_{u_0}^2 + tk\sigma_\delta^2 + k\sigma_\eta^2, \text{if } k < t, \\
\text{var}(r_{it}) - \text{var}(a_{it}) = \sigma_{u_0}^2 + t^2\sigma_\delta^2 + t\sigma_\eta^2 + \sigma_\eta^2
\]

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

References


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


Table 1: Summary Statistics, SIPP 1996

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

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

Table 2: Total number of observed job changes (in percentages) and within-job wage growth, by experience

<table>
<thead>
<tr>
<th>Quartiles of initial labor market experience (period)</th>
<th>Number of job changes</th>
<th>Within-job Δw</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Less than 25th (0-2)</td>
<td>31.2</td>
<td>29.3</td>
</tr>
<tr>
<td>25-50 (3-6)</td>
<td>46.8</td>
<td>29.9</td>
</tr>
<tr>
<td>50-75 (7-13)</td>
<td>60.0</td>
<td>23.5</td>
</tr>
<tr>
<td>More than 75th (&gt;13)</td>
<td>72.1</td>
<td>18.9</td>
</tr>
<tr>
<td>Total</td>
<td>52.2</td>
<td>25.5</td>
</tr>
</tbody>
</table>
Table 3: Tobit regression of total number of job changes

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Married</td>
<td>-0.11</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Own a house</td>
<td>0.55***</td>
<td>(0.20)</td>
</tr>
<tr>
<td>White</td>
<td>0.08</td>
<td>(0.14)</td>
</tr>
<tr>
<td>High school</td>
<td>-0.06</td>
<td>(0.22)</td>
</tr>
<tr>
<td>At least some college</td>
<td>-0.21</td>
<td>(0.21)</td>
</tr>
<tr>
<td>Initial experience level</td>
<td>-0.10***</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.93***</td>
<td>(0.21)</td>
</tr>
<tr>
<td>Observations</td>
<td>1211</td>
<td></td>
</tr>
</tbody>
</table>

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 own 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

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

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

Note: Probit regressions of job mobility on lagged within job wage growth. The dependent variable is 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, two- and three-period lagged level of wage. Standard errors (in parenthesis) are clustered by person.
Table 6: Estimated model parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>With job-switching cost</th>
<th>Ignoring job-switching cost</th>
<th>Ignoring match-level shocks</th>
<th>Neglecting firm-specific wage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Wage shocks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma^2_\eta \times 100 )</td>
<td>0.829</td>
<td>0.624</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.020)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma^2_\zeta \times 100 )</td>
<td>0.005</td>
<td>0.003</td>
<td>0.176</td>
<td>0.321</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.033)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>( \sigma^2_v )</td>
<td>0.032</td>
<td>0.037</td>
<td>0.036</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Random growth factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu_c )</td>
<td>-0.005</td>
<td>-0.006</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma^2_c \times 10000 )</td>
<td>0.108</td>
<td>0.496</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.226)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu_\delta )</td>
<td>0.012</td>
<td>0.004</td>
<td>-0.002</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>( \sigma^2_\delta \times 10000 )</td>
<td>0.255</td>
<td>0.010</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.008)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma^2_{a_0} )</td>
<td>0.003</td>
<td>0.015</td>
<td>0.065</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.010)</td>
<td></td>
</tr>
<tr>
<td>( \sigma^2_{u_0} )</td>
<td>0.068</td>
<td>0.084</td>
<td>0.072</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.002)</td>
<td>(0.005)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Other labor market parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \lambda^e )</td>
<td>0.483</td>
<td>0.849</td>
<td>0.730</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.053)</td>
<td>(0.022)</td>
<td></td>
</tr>
<tr>
<td>( \gamma_{constant} )</td>
<td>0.698</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma_{married} )</td>
<td>-0.377</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma_{house} )</td>
<td>-0.290</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma_{hs} )</td>
<td>0.086</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma_{college} )</td>
<td>-0.386</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma_{white} )</td>
<td>-0.277</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Bootstrap standard errors are in parentheses. \( \sigma^2_\eta \), \( \sigma^2_\zeta \) and \( \sigma^2_v \) are, respectively, the variance of match- and person-level shock, and transitory shock (measurement error). \( c \) and \( \delta \) are the random growth factor at match and person level respectively. \( \sigma^2_{a_0} \) is the heterogeneity in the match values of job offers. \( \sigma^2_{u_0} \) is the heterogeneity in the person-component of wages at the start of work life. \( \lambda^e \) is the offer arrival probability. \( \gamma \)'s are switching cost parameters.
Table 7: Life-time insurance value of job mobility

<table>
<thead>
<tr>
<th>Switching Cost Scenario</th>
<th>Mean $\Delta V_0$</th>
<th>$\Delta V_0/V_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Married, do not own a house, college graduate, white</td>
<td>-0.34</td>
<td>2.21</td>
</tr>
<tr>
<td>Single, own a house, college graduate, white</td>
<td>-0.26</td>
<td>2.17</td>
</tr>
<tr>
<td>Married, do not own a house, high school graduate, white</td>
<td>0.13</td>
<td>1.78</td>
</tr>
<tr>
<td>Married, do not own a house, high school graduate, black</td>
<td>0.41</td>
<td>1.56</td>
</tr>
<tr>
<td>Single, do not own a house, less than high school, white</td>
<td>0.42</td>
<td>1.55</td>
</tr>
</tbody>
</table>

Figure 1: Process of job mobility decisions

Figure 2: Match component and job change
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 $\sigma_\eta^2$ and $u_c$ on within-job wage growth over life $E(a_{t+1} - a_t | M_{t+1} = 0)$

Note: Simulated from the model without switching cost. Baseline: $u_c = -0.006$, $\sigma_\eta^2 = 0.0062$
Figure 5: The effect of $\sigma^2_{w}$ and $\sigma^2_\eta$ on the autocovariance of job mobility over life

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

Figure 6: Decomposing the experience-profile of 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.
Figure 7: Changes in the variance of realized match value over time, for a given type of worker

![Graph showing changes in variance over time](image)

Figure 8: Insurance value of on-the-job search at the beginning of life, for different switching costs $k_3 > k_2 > k_1$

![Graph showing insurance value vs. $a_0$](image)