# Cross-Sectional Volatility of Individual Income Processes*

Damba Lkhagvasuren†
Concordia University

September 2009

There is large variation in individuals' earnings within narrowly-defined local labor markets even after controlling for age and education. Moreover, a considerable part of this variation is due to transitory labor income shocks. However, each year a small fraction of the labor force moves between different markets. Given these facts, estimation of structural models that assume that the transitory earnings component is responsible for gross labor mobility yields very high moving costs. In this paper, I demonstrate that ignoring the positive correlation of labor income shocks across space, when in fact it is present, introduces a large upward bias into the estimates of moving costs. First, I construct a multi-factor dynamic Roy's model of inter-sectoral mobility by explicitly allowing for: 1) both net and gross mobility, and 2) correlation of transitory earnings across space. Second, I estimate the model using micro-data on individual level mobility patterns and earnings differences between movers and stayers. The model makes a strong prediction that within each age-and-education group, individuals with lower unobserved ability have lower moving costs. It also indicates that the cross-sectional variation of idiosyncratic labor income shocks of college educated workers is more than twice that of high school educated workers. These findings provide new insights into the mobility differences across schooling levels and therefore cast doubt on the role of a credit constraint in explaining the observed differences.

**Keywords:** Regional Mobility, Labor Income Processes, Roy's Model, Moving Cost, Skill Composition of Movers versus Stayers

*This is a revised version of chapter three of my Ph.D. dissertation submitted to University of Rochester. I owe many thanks to my advisor Mark Bils. I benefited greatly from conversations with Árpád Ábrahám, Marine Carrasco, Kevin Clarke, Jonas Fisher, Fatih Guvenen, Lutz Hendricks, Chris Taber, Michael Raith and the seminar participants at University of Rochester, Concordia University, the 2007 Midwest Macro Meetings and the 2009 Canadian Economic Association Meetings. All remaining errors are my own.

†Contact address: Department of Economics, Concordia University, 1455 Maisonneuve Blvd. W, Montreal, QC H3G 1M8; e-mail: dlkhagva@alcor.concordia.ca

# 1 Introduction

There is large variation in individuals' earnings within narrowly-defined labor markets even after controlling for age and education. Moreover, a considerable part of this variation is due to transitory labor income shocks. However, each year a small fraction of the labor force moves between different local markets. Given these facts, estimation of structural models that assume that the transitory income shocks are responsible for gross labor mobility and that these shocks are uncorrelated across local labor markets yields very high moving costs. To make the point clearer, let us consider the following simple thought experiment. If the idiosyncratic location-specific shocks are highly correlated, the expected gain of switching locations would not be substantial and therefore there will be very low mobility even when the moving cost is very small. Now, if we make the location-specific shocks are less correlated, more workers would move between locations. Therefore, low mobility and high within-market variance of transitory earnings should not necessarily imply high moving costs.

In this paper, I argue that the moving cost and the correlation of idiosyncratic labor income shocks across sectors are intimately related. To study the implications of this relation, I construct a two-sector model in which regional mobility and earnings' differences between movers and stayers are endogenously determined. I estimate the model using micro data on individuals' income and regional mobility for two main schooling levels, college and high school graduates.[1] Specifically, I focus on the following key data features documented in this paper:

1. Wage gap between movers and stayers[2]
   (a) Among high school educated workers, migrants earn less than non-migrants at the receiving localities;

[1] In this paper, I abstract from schooling decisions and therefore, look at individuals who have already attained their highest schooling whether it is a high school diploma or a college degree. This assumption, although restrictive, allows me to reduce the number of state variables and makes the model computationally more tractable. For the studies of the interaction of schooling and geographic mobility, see, for example, Heckman et al. (1996) and Dahl (2002).

[2] Wage gap between movers and stayers refers to earnings differences between movers and stayers in the destination region. In other words, it measures earnings of newly-arrived workers relative to earnings of local residents within a given region.

(b) College-educated movers earn more than their non-migrant counterparts;

(c) For high school educated workers, the wage gap between movers and stayers decreases with age;

(d) For college educated workers, the gap increases over the life-cycle;

2. Net and gross mobility

(a) Gross mobility increases with schooling level: college educated workers are more mobile than high school graduates; and

(b) Net mobility is higher for less educated workers: less-educated workers are relatively more responsive to local labor market conditions.

I show that ignoring the positive correlation of labor income shocks across sectors, when in fact it is present, introduces large biases in the structural parameters of a multi-sector model. The estimated moving costs are about 11-16% of average annual earnings in the benchmark model where the cross-correlation of the shocks is allowed, compared to 60-80% of average annual earnings in the restricted models where the shocks are uncorrelated. The assumption that the idiosyncratic transitory shocks are not correlated across regions has counterfactual implications on earnings differences between movers and stayers. It also requires implausibly large regional shocks to generate the observed net mobility.

Using the model, I address several issues relating to regional mobility: Does the lower mobility rate of less-educated workers imply that they have higher moving costs relative to their labor income? Are credit constraints responsible for the positive relation between schooling and mobility?[3] What are the factors responsible for the measured differences between earnings of movers versus stayers? Who moves and who stays behind by unobserved skill levels?

The estimation of the model reveals that approximately 10% of the transitory earnings variation among high school educated individuals is due to the pure location-match effect. However, for college graduates, the location-match effect amounts to 25% of the transitory earnings variation. More interestingly, the moving costs relative to labor income do not

[3]In current literature, the lower mobility of less-educated workers is mainly explained by credit constraints or high moving costs relative to their wage. For example, see Wozniak (2006) for discussions of observationally equivalent theories of educational differences in migration that assume that workers with lower education are credit constrained.

differ much between the two education groups. Therefore, college educated workers move more, not because they have lower relative moving costs, but because their transitory earnings shocks are more volatile across locations than those of high school educated workers. These findings suggest that market thickness decreases with education: for college graduates different markets offer different opportunities whereas for less-educated workers, employment opportunities do not vary much across different local markets.

I find that, for high school educated workers, newly-arrived workers earn less than local residents because workers with lower unobserved ability have lower moving costs and therefore move more frequently than those with higher ability. However, for college educated workers, the effect of this permanent ability is offset by the effect of location-specific transitory shocks. In other words, the positive sorting along the transitory shocks is much stronger for college educated workers due to their cross-sectionally more volatile earnings shocks. These results cast doubt on the role of a credit constraint in explaining mobility differences across schooling levels.

There is an extensive literature which discusses the characteristics of movers by observable skill levels such as age and education (Greenwood, 1997). However, little is known about who moves and who stays behind within each age and education group. The model predicts that, within each age and education group, individuals with lower unobserved ability have lower moving costs and move more frequently than those with higher unobserved ability.

The importance of estimating structural parameters of regional mobility models goes far beyond understanding local labor market dynamics. We are ultimately interested in quantifying responses to both aggregate and local policy interventions, as well as their welfare implications. For example, without a reasonable estimate of the magnitude of the moving costs, it would be difficult to interpret the evaluation of competing government policies that influence individuals' job search across sectors. Positively speaking, estimation of structural parameters in a multi-sector model can also help understand how the differences in the ease of switching local markets, whether viewed as search costs or training costs, affect cross-country differences in aggregate variables such as labor supply

and wage.[4]

There are several factors responsible for the results of this paper. First, I build on Roy's model while considering a richer structure on individuals' income processes. In his 1951 paper, Roy outlines a simple model of selection based on comparative advantage. He investigates the resulting effects on the distribution of earnings in different occupations. Roys framework has been applied to a number of inter-sectoral mobility studies, including internal and international migration (Nakosteen and Zimmer, 1980; Borjas, 1987; Borjas, 1992; Dahl, 2002) and inter-industry mobility (Heckman and Sedlacek, 1990, McLaughlin and Bils, 2001). Most of the existing regional mobility studies that use the conceptual framework of Roy's model share the same feature: agents make only one decision, namely, which market or sector to enter? My departure from these studies is that I consider a dynamic setting where individuals make reversible mobility decisions.

One paper that allows for reversible mobility is Kennan and Walker (2006), which develops a fully specified structural migration model with many alternative locations and estimates it using panel data on young high school graduates. Their model assumes a stationary environment for the local markets and hence considers only gross mobility. My findings suggest that the relative magnitude of net and gross mobility is key to measuring the cross correlation of local labor income shocks along with other important structural parameters. Moreover, Kennan and Walker make an informational assumption that, in order to determine the wage in another location, individuals have to move, whereas in my model agents know their wage in a new location even before they move. Put differently, their model assumes undirected search across labor markets while ours use directed search. Directed search is essential for my model, since I directly compare the wage differences between movers and stayers.

The second important factor responsible for the predictions of the model is wage differences between movers and stayers. Measuring the wage differences between movers and stayers is an important empirical analysis in its own right. Over the past decades, much has been made of regressing wage on an individual's migration status in addition to

[4]See, for example, Lkhagvasuren (2009) for a discussion about different channels through which moving cost affects aggregate labor market equilibrium.

other individual characteristics.[5] Most of the empirical work in this literature has focused on assessing the returns to migration. These studies find that migrants tend to have lower earnings than non-migrants in the receiving localities. It is interesting to notice that empirical studies which do not explicitly focus on young workers, tend eventually to end up focusing on them. This is because of the interaction of their highly restrictive parametric assumption on the age-earnings profile and the non-monotonic relation of age and mobility. To measure the earnings differences between movers and stayers, first, I maintain a less restrictive parametric assumption on the age-earnings profile. Second, to correct for the bias in the measured differences due to cross-sectional variation in earnings, I nest the U.S. Census with the Annual Demographic Survey (March CPS) by a rank statistic. It should be mentioned that the observed wage patterns are absent in commonly-used panel data sets as they are not well suited to use for regional mobility as each year a small fraction of their sample moves across regions.

Most structural studies on individual migration decisions tend to focus on young workers. This is because migration propensities are much higher among young workers and narrowing the analysis on a less heterogeneous demographic group aids computational tractability of structural models. Thus, there is little empirical evidence on the relative importance of the individual life-cycle effect on internal migration. As Greenwood (1997) notes: *"The influences of life-cycle changes on migration decisions have only barely been touched by researchers."* I measure the relative moving cost of old workers using micro-data on wage dynamics of old movers. The model with different age groups aids the identification of the selection effect by comparing movers' earnings across age groups.

The outline for the rest of the paper is as follows. In Section 2, I present the main empirical findings on earnings differences between migrant and non-migrant workers in the United States, along with the key facts of inter-regional mobility across age-and-education groups. In Section 3, I construct a dynamic multi-factor Roy model in a directed search setting while explicitly allowing for both gross and net mobility. In Section 4, I estimate the model using the empirical moments measured in Section 2, and in Section 5, I draw together the conclusions of the paper.

[5] Greenwood (1997) presents an extensive survey of the literature.

# 2 Facts

The key variables in my analysis are mobility and the wage rate. For the reasons explained below, my main geographic unit will be Census Divisions. CPS records the sub-national geographic units where households resided a year ago, whereas the U.S. Census asks respondents which state the individuals lived in five years ago. Both of the data sources ask individuals' current state of residence. Therefore, to construct the individual's mobility status, I use the states of residences at two consecutive time periods. For the wage rate, I use the ratio of annual labor income and annual total hours worked.

My empirical analysis focuses on white male workers between the ages of 20 and 64 years with the education levels of high school or college. My sample includes individuals who has worked more than 750 hours and less than 5000 hours per year and has hourly earnings more than \$2.00 and less than \$400.00 with 1993 as the base year. I distinguish two age groups: *young*, aged 20-34, and *old*, aged 35-64. In the remainder of the paper, I let $a$ and $c$ denote the actual yearly age and the age group, respectively: $a \in \{20, 21, ..., 64\}$ and $c \in \{young, old\}$.

## 2.1 Mobility by age and education

Table 1 shows gross labor mobility between Census-Divisions. Gross mobility is highest among young college graduates and lowest for old high school educated workers. Although these differences in gross mobility across age and education groups are well known in the literature, it is not immediately clear how mobility differs across local labor markets. Using the U.S. data, Lkhagvasuren (2009) shows that most of the moves across local labor markets cancel out: both out- and in-migration are much larger than their differences at the local level. The extent that in- and out-migration cancel out at the local level provides an important insight on the determinants of labor mobility. For example, if out- and in-migration of the observationally same individuals fully cancel out, it would suggest that the idiosyncratic shocks are primarily responsible for labor mobility. On the contrary, if the two flows differ from each other substantially, it would indicate that local labor market conditions or sectoral shocks are responsible for the bulk of labor mobility.

Table 1: Mobility by Age and Education: Census 1980-2000

| *Education* | *High School* | | *College* | |
|---|---|---|---|---|
| | young | old | young | old |
| Mobility Rate | 0.089 | 0.036 | 0.171 | 0.072 |
| | (0.0093)* | (0.0039) | (0.0022) | (0.0074) |
| Volatility of out-migration | 0.0205 | | 0.0145 | |

*Standard errors are in parenthesis.

To examine whether there is a substantial difference between education groups in the extent that in- and out-migration cancels out, we look at the cross-sectional volatility of out-migration. Table 1 reports the measured standard deviations of out-migration for each education group. For both groups, the measured standard deviations are much lower than the mean migration rate implying that gross mobility is much larger than net mobility. What is more remarkable about these estimates is that high school graduates move more with net mobility relative to college graduates.[6] In addition to these differences across age and education groups, mobility can also differ by unobserved ability. If this is the case, one would expect significant earnings differences between movers and stayers within each age and education cell. In the remainder of the section, I measure these differences while controlling for age and education.

## 2.2 Earnings differences between movers and stayers

To measure the unobserved skill differences between movers and stayers, I follow McLaughlin and Bils (2001), and compare the earnings between movers and stayers within the same locality. Consider an individual $i$ who is working in location $j$ at time $t$ and whose age

[6] I do not look at the relative magnitude of net mobility by age groups. This is because of a measurement issue. Regardless of the size of the data set, there are very few old movers at the sub-national level for each year. Therefore, small fluctuations in their numbers generate high variance of mobility across regions. This might give a false indication that old workers are highly sensitive to local labor market conditions. Nevertheless, Lkhagvasuren (2006) examines contemporaneous correlations between in-migration and out-migration across states for different age groups and finds that the two flows are negatively correlated for younger workers and uncorrelated for older workers. Therefore, older workers are less susceptible to local conditions than younger workers.

and schooling level are $a$ and $s$, respectively. Suppose that his wage is

$$w_{ijts}(a) = G_s(a) + \eta_{jts} + \mu_i + \epsilon_{ijts} \tag{1}$$

where $G_s(a)$ is a known age-earnings profile for education level $s$, $\eta_{jts}$ is the regional effect, $\mu_i$ is an individual-specific effect and $\epsilon_{ijt}$ is the transient effect. In the above equation, the aggregate effect can be thought of as an interaction of the time effect and the regional effect, as I do not restrict $E_j\eta_{jts}$ to be zero.[7]

Let $t_1$ and $t_2$ denote two time periods such that $t_1 < t_2$. If a person resides in region $j$ during $t_1$ and in region $j'$ during $t_2$, we call him a newly-arrived of $j'$ at $t_2$ with respect to $t_1$ and set $d(t_1, t_2, j') = new$. If a person works in region $j$ for both periods, we call him stayer of $j$ at $t_2$ (with respect to $t_1$) and set $d(t_1, t_2, j) = local$. Now, given two time periods $t - \Delta$ and $t$, and region $j$, the mean wage difference between stayers and newly-arrived workers is given by

$$\begin{aligned} D_{jts}(\Delta, a) &= E_i(w_{ijts}|a, d(t-\Delta, t, j) = local) - \\ &\quad -E_i(w_{ijts}|a, d(t-\Delta, t, j) = new) \\ &= E(\mu_i + \epsilon_{ijts}|a, d(t-\Delta, t, j) = local) - \\ &\quad -E(\mu_i + \epsilon_{ijts}|a, d(t-\Delta, t, j) = new). \end{aligned} \tag{2}$$

Several issues arise in measuring this selection effect. McLaughlin and Bils (2001) point out that data for estimating earnings differences between migrant and non-migrant workers, $D_s(\Delta, a) = E_{jt}\big(D_{jts}(\Delta, a)\big)$, must be large panel data, as movers are a small fraction of the population. Although PSID consistently records individual geographic information such as state and county of residence, its small sample size makes it impossible to construct meaningful empirical patterns at sub-national levels. Therefore, the first problem is availability of large panel data of a representative sample of the working age population. To make the point clear, let us consider $n$ regions. Let $P(n)$ be the average number of people residing in a region and let $m(n)$ denote the mobility rate. Then the lower bound of the cross-sectional variance of the measured differences will be

$$\mathrm{var}_x(D) \geq \frac{\mathrm{var}_x(wage)}{P(n)m(n)}.$$

[7] $E_y$ denotes the expectation taken over $y$.

Therefore, in order to reduce the variance, one has to go with larger sub-national units, $P(n)$. However, migration is known to decrease as distance increases. In other words, the mobility rate, $m(n)$, decreases with the size of the local labor markets, $P(n)$. Therefore, one has to find a balance between the size of the region and the size of the sample. Recognizing these issues, I proceed with the U.S. Census and CPS, taking the nine Census Divisions as my basic geographic units.

The second issue is the interaction of the low number of movers and earnings variance. Although my empirical focus is the mean earnings of movers, within-region earnings variance[8] can make the measured mean earnings highly volatile when the sample size is small. Individuals from the same skill level but from different regions will unevenly contribute to the variance of movers' wage, var$(D(\Delta))$. On the other hand, we know that the earnings variance increases over the individual life-cycle. Given very few observations, the residual earnings of a few old workers can also explode the variance of the estimated mean residual earnings for movers.

To take into account this issue, I use a rank statistic instead of the actual earnings measured in dollars. Specifically, given time $t$ and earnings profile $G_s(a)$, I order individuals in the same region by their residual earnings, $w_{ijts}(a) - G_s(a)$, and determine their relative rank, $r_{ijts}(a)$, in the within-region distribution:

$$r_{ijts}(a) = \frac{\left(\sum_k I_{w_{ijts}(a)-G_s(a) \geq w_{kjts}(a)-G_s(a)}\right) - 1}{N_{jts} - 1}$$

where $N_{jts}$ is the number of workers who have schooling level $s$ and resides in region $j$ at time $t$ and $I$ denotes the index function. As the ranking is carried out within each region, it follows that

$$r_{ijts}(a) = \frac{\left(\sum_k I_{\mu_i+\epsilon_{ijts} \geq \mu_k+\epsilon_{kjts}}\right) - 1}{N_{jts} - 1}.$$

Further, given that the number of stayers is much larger than the number of movers, $E(r_{ijt}|\text{local})$ will be very close to 0.5. Therefore, given the interval between the two time points $\Delta$, I focus on the mean rank of movers at their destination for each age and education group:

$$\tilde{D}^r_s(\Delta|c, new) = E_{iajt}(r_{ijts}(a)|a \in c, new) \tag{3}$$

[8]See Gomme and Rupert (2005) for discussion on wide cross-state dispersion of income in the U.S.. As each Census Division is comprised from several states, there is also substantial within-division dispersion.

where $c \in \{\text{young, old}\}$ and $s \in \{\text{high school, college}\}$.

Finally, there is an issue of the parametric assumption on age earnings profile. If we do not allow for highly flexible function for $G$ then there will be a certain age $a'$ where $E(\mu_i + \epsilon_{ijts}|a') \neq 0$. In other words, there will be a nonzero function $g(a) = E(\mu + \epsilon)$. Therefore, unless the non-monotonic relation between age and mobility, $m(a)$, is perfectly aligned with the error function, $g(a)$, the measured differences, $D(\Delta)$, will be affected by the miss-specification of the age-earnings profile. To handle this issue, I use a quartic polynomial for $G$ and estimate the coefficients using OLS. Assuming a higher degree polynomial (5-th or 6-th order) for $G$ does not change the result.

### 2.2.1 Cross-validation

To cross-validate the specification, I bring the model to the Census data and compare the results with the mean wage differences of movers and stayers for *each single* age and education cell. Instead of focusing on the broader age groups $c \in \{young, old\}$ as we have done so far, let us narrow down to each single age $a$, $a \in \{20, 21, \cdots, 64\}$ while fixing the schooling level. For region $j$, period $t$, and schooling level $s$, let us consider the following rank statistic within each age:

$$r^w_{ijts}(a) = \frac{(\sum_k I_{w_{ijts} \geq w_{kjts}}) - 1}{N_{jts} - 1}$$

for each $a$. Notice that the rank is constructed using the actual wage $w_{ijts}$ as opposed to the residual wage $\mu_i + e_{ijt}$. Let the associated conditional mean rank be

$$\bar{r}^w_{jts}(a|new) = E_i(r^w_{ijts}(a)|\ new).$$

Then, a natural parallel of $\tilde{D}^r_s$ will be

$$\tilde{D}^w_s(\Delta|c,\ new) = E_{ajt}(\bar{r}^w_{jts}(a|a \in c,\ new))$$

for each $c \in \{\text{young,old}\}$ and $s \in \{\text{high school, college}\}$. By construction, $\tilde{D}^w$ measures the relative performance of newcomers without any parametric assumption on earnings profile $G$.

Table 2: Mean Wage Rank of Movers

| *Mean Rank* | *High School* | | *College* | | *Data Source* |
|---|---|---|---|---|---|
| | young | old | young | old | |
| $\tilde{D}^r(\Delta = 1)$ | 0.471* | 0.459 | 0.522 | 0.542 | CPS |
| $\tilde{D}^r(\Delta = 5)$ | 0.443 | 0.474 | 0.517 | 0.541 | Census |
| $\tilde{D}^w(\Delta = 5)$ | 0.443 | 0.475 | 0.516 | 0.542 | Census |
| $\tilde{D}^r(1 < \tau_m < 5)$ | 0.434 | 0.478 | 0.514 | 0.542 | CPS and Census |

### 2.2.2 Measured differences

Table 2 displays the estimates of (3) which were obtained using Census and CPS. $D(\Delta)$ denotes the mean rank difference between local residents and newly-arrived workers who moved in during the last $\Delta$ periods. If there were no earnings difference between movers and stayers these moments would be equal to one half. However, if movers earn less (more) than the local residents, $D(\Delta)$ will be lower (higher) than one half. Given the measured rank statistics, Table 2 documents four important features of data which we presented in the introduction. Next, I look at the wage differences between movers and stayers within each yearly age and education cell for a cross-validation purpose. Table 2 shows that the earnings gap measured using $\tilde{D}^w$ in Censuses 1980 to 2000 is highly consistent with the one obtained using $\tilde{D}^r$. Moreover, the results from the U.S. Census is consistent with those from BLS's Current Population Survey despite the smaller sample size of the latter.

### 2.2.3 Correcting for data timing

As we discussed earlier, in both the U.S. Census and CPS, the only consistent variable on interregional mobility are the region where a household resided a certain time period ago and where it is residing now. More specifically, given the two data sources, it is impossible to know when the relocation actually took place. For example, in the Census, it is not clear when exactly during the last five years the person moved into their new localities. Similarly, we can not tell which month during the past twelve months the CPS households switched their residences. Therefore, the measured wages for movers are not the actual wage rates earned at their new location. For example, consider an individual who worked in region $A$ during the first nine months of the previous calender year and worked in region $B$ for the remainder of that year. Suppose that his wage rates were \$10.00 per hour and \$12.00 per hour in $A$ and $B$, respectively. The reported annual hourly wage rate for the newly arrived worker will be $\$\frac{9*10.00+3*12.00}{12} = \$10.50$ which is much lower than \$12.00/hour, the actual wage rate in the new location.

If most of the mobilities occur at the beginning of the year then we would not have much problem as the bulk of the data collection occurs in the first half of the year for the two data sources. However, it is well known that inter-regional mobility peaks during the summer time. Therefore, without taking into account this serious bias, it is impossible to proceed with the wage rates directly measured from these main data sets.

To correct for the bias, I nest the mean ranks measured from the Census with those measured from the CPS. Let $\tau_m$ denote the uninterrupted time period an individual has worked in his current location. Let $\tilde{D}^r(\tau_m = x)$ and $P(\tau_m = x)$ denote the mean wage rank and the number of workers who arrived $x$ periods ago respectively. Then we will have

$$\frac{P(\tau_m < 1)\tilde{D}^r(\tau_m < 1) + P(1 < \tau_m < 5)\tilde{D}^r(1 < \tau_m < 5)}{P(\tau_m < 1) + P(1 < \tau_m < 5)} = \tilde{D}^r(\tau_m < 5).$$

Therefore the unbiased estimate for the earnings difference between movers and stayers is given by

$$\tilde{D}^r(1 < \tau_m < 5) = \frac{P(\tau_m < 5)\tilde{D}^r(\tau_m < 5) - P(\tau_m < 1)\tilde{D}^r(\tau_m < 1)}{P(\tau_m < 5) - P(\tau_m < 1)}$$

Notice that $\tilde{D}^r(\tau_m < 1) = \tilde{D}^r(\Delta = 1|new) = \tilde{D}^r_{CPS}$ and $\tilde{D}^r(\tau_m < 5) = \tilde{D}^r(\Delta = 5|new) =$

$\tilde{D}^r_{CEN}$ where subscripts denote the data sets the moments are measured from. Therefore, the corrected earnings gap between movers and stayers is given by

$$\tilde{D}^r(1 < \tau_m < 5) = \frac{P_{CEN}\tilde{D}^r_{CEN} - P_{CPS}\tilde{D}^r_{CPS}}{P_{CEN} - P_{CPS}}.$$

Table 2 reports the estimates of $\tilde{D}^r(1 < \tau_m < 5)$ for all four groups.

# 3 Model

## 3.1 Preliminaries

My goal is to construct and estimate a model of regional mobility that is flexible enough to take into account the facts presented above. I begin with the static setting of the standard Roy's self-selection model and extend it into a dynamic setting. Consider two locations denoted by $A$ and $B$. Individuals differ by their productivity that can differ across location. A worker is characterized by a pair of the location-specific productivity levels denoted by $(\varphi_A, \varphi_B)$. I decompose the location-specific productivity into a permanent and a transitory component. I assume that the permanent component is common across the two locations, i.e. for any $j \in \{A, B\}$,

$$\varphi_j = \mu + e_j$$

where $\mu$ and $e_j$ denote the permanent and transitory components, respectively.

Assuming that there are two levels of the permanent component, $\mu_L$ and $\mu_H$, the productivity distribution of the heterogeneous workers can be presented as in Figure 1. Each of the large number of heterogeneous workers is described by a point on the graph. The iso-probability contours denoted by $\mu_p$ reflect the productivity distribution of workers with the permanent component of $\mu_p$ for both $p \in \{L, H\}$.

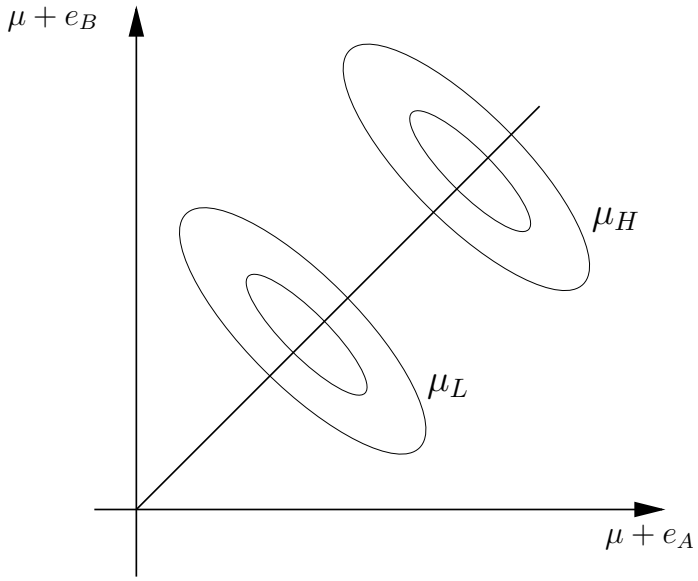
Let $z_j$ denote the location-specific technology shock which is common to all workers in location $j \in \{A, B\}$. If the person works in location $A$, her wage will be

$$w_A = z_A + \varphi_A.$$

Similarly, if the person works in location $B$, her wage will be

$$w_B = z_B + \varphi_B.$$

Figure 1: Dynamic Roy Model with Zero Moving Cost




Each worker is described by a point in the graph. The iso-probability contours reflect the distribution of the transitory productivity components $(e_A, e_B)$ of within each permanent ability levels: $\mu_L$ and $\mu_H$. When the moving cost is zero and $z_A = z_B$, the individuals who are indifferent between working in location $A$ and location $B$ lie along the bisectrix.

Let us assume that individuals draw their utility only from their wage. Also, let us assume for now that there is no cost associated with switching the two locations. Then, there can be a set of workers who are indifferent between working in $A$ and $B$. The set is a line given by $e_B = (z_A - z_B) + e_A$ which we refer to as an indifference line. When the location-specific shocks are equal, $z_A = z_B$, the indifference line overlaps with the $45^\circ$ line (see Figure 1).

Let us turn off the location-specific technology shocks, $z_A = z_B = 0$. Then, due to the transitory shock $(e_A, e_B)$, individuals move in both directions between the two islands. Now, if the location $A$ is hit by a small positive permanent technology shock, $z_A > 0$, the indifference line will shift up creating a gap between the two mobility flows: the number of workers moving from $A$ to $B$ will be smaller than the number of workers moving in the opposite direction. Whether the average wage of newcomers is higher or lower than the

average wage of local residents of $A$ will depend both on the shape of the productivity distribution[9] and the magnitude of the regional shock, $z_A$. Therefore, this simple model has an important implication that the overall earnings gap between newly-arrived workers and local residents are related to the magnitude of net mobility relative to gross mobility. Another important observation is that the sensitivity of local labor force to local labor market condition is positively related to the cross-correlation of the idiosyncratic shocks. For example, in the extreme case where the idiosyncratic shocks are perfectly correlated, individuals move with only net mobility.

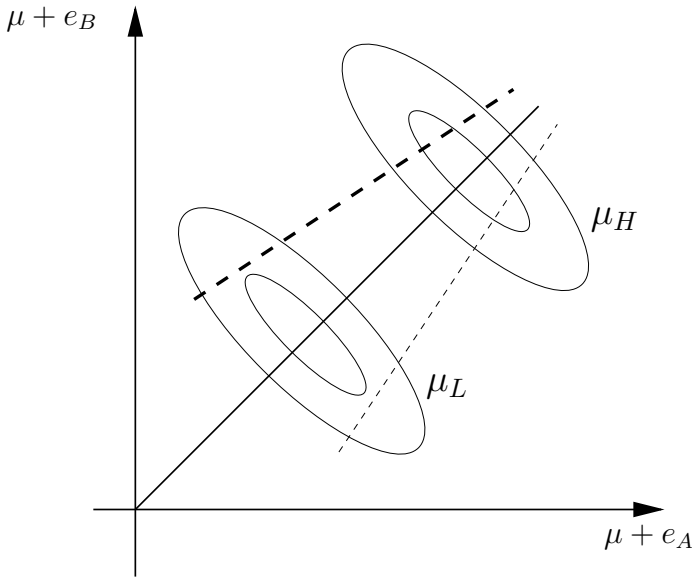
If we introduce a non-zero moving cost, $T$, it will also affect the differences in earnings between movers and stayers. Further, if we allow the moving cost $T$ to differ by the individual fixed effect $\mu$ (See Figure 2), the earnings differences between movers and stayers will depend on two additional factors, the mobility differences between workers with different permanent income components $\mu$, and the relative magnitude of transitory versus permanent components of individual labor income processes.

Maintaining more flexible shape for the location-specific productivity shocks comes with the cost of a higher dimension of the state space in the stochastic dynamic model. Moreover, the number of state variables associated with the individuals decision problem increases more than proportionally with the number of locations. Therefore, given other important heterogeneity such as age, education and unobserved ability, I proceed with a two-sector Roy's model. The two sectors can be thought of as an approximation to the interaction of local labor market and the rest of the economy. Nevertheless, the qualitative conclusions of the paper will not change if we introduce more locations into the model.

Finally, I incorporate the individual life-cycle effect into my model by allowing the moving costs to differ across different age groups. Having different age groups aids the identification of the model by comparing earnings of movers with different age groups while fixing the education level. In the remainder of the paper, I formalize these ideas and estimate the parameters of the model for the two different schooling levels.

[9]See McLaughlin and Bils (2001) for an explicit discussion of how the shape of the distribution affects mover-stayer wage comparison in the standard Roy's model. However, certain caution should be taken in comparing their work and this paper as the former allows only for net mobility whereas this paper allows for both net and gross mobility.

Figure 2: Dynamic Roy Model with Non-zero Moving Cost




Let $z_A = z_B$. When there is a nonzero moving cost, workers who are indifferent between working in $A$ and moving to $B$ are aligned along the heavy dashed line. Similarly, those who are indifferent between staying in $B$ and moving to $A$ are aligned along the light dashed line.

## 3.2 Detailed model specification

### 3.2.1 Environment

The economy consists of two islands denoted by $A$ and $B$. The islands are inhabited by risk-neutral workers which could be young or old. Young agents get older each period with probability $\delta^y$. Old agents retire with probability $\delta^o$ and are replaced by newly-born workers. Individuals have one of the two schooling levels: high school and college. Let us denote the schooling level by $s$ with $s = 1$ for high school educated workers and $s = 2$ for college educated workers. For the remainder of the paper, I refer to individuals from the same age and schooling level as a group.

The two locations require different sets of skills. In addition to age and schooling levels, individuals are characterized by their permanent, unobserved ability $\mu$ which does not change over their life-cycle and across locations. For $\mu$, I assume a zero-mean normal

distribution with variance $\sigma_\mu^2$, $\mu \sim N(0, \sigma_\mu^2)$. Individuals are also subject to idiosyncratic location-specific transitory shocks $(e_A, e_B)$. Both $e_A$ and $e_B$ follow the same AR(1) process:

$$e_j' = \rho_e e_j + u_j$$

where $j \in \{A, B\}$ and location-specific innovations $(u_A, u_B)$ are drawn from a bivariate normal distribution

$$\begin{pmatrix} u_A \\ u_B \end{pmatrix} = N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_u^2 & \rho_x \sigma_u^2 \\ \rho_x \sigma_u^2 & \sigma_u^2 \end{pmatrix}\right)$$

where $\rho_x$ denotes the correlation between the two location-specific innovations, $u_A$ and $u_B$.

Each island is subject to a local income shock $z_j$, $j \in \{A, B\}$, which also follows an AR(1) process: $z_j' = \rho_z z_j + v_j$ where $v_j \sim N(0, \sigma_v^2)$. So, the local income shocks are not correlated across islands. Let $\sigma_z^2$ denote the variance of the regional shock, $z$. Each period, individuals choose to work on one of the two islands. Individuals move between the islands subject to a moving cost, $T(c, s, \mu)$, which could depend on the individual's age, $c$, schooling, $s$, and the unobserved ability level, $\mu$.

### 3.2.2 Wage

Both labor markets are competitive and individuals are paid their marginal products. Individuals are fully characterized by $(c, s, \mu, e_A, e_B)$, and their current residence, $j$. For a worker who is working on island $A$, the wage will be

$$w_A(z_A, z_B, e_A, e_B, \mu) = \mu + z_A + e_A.$$

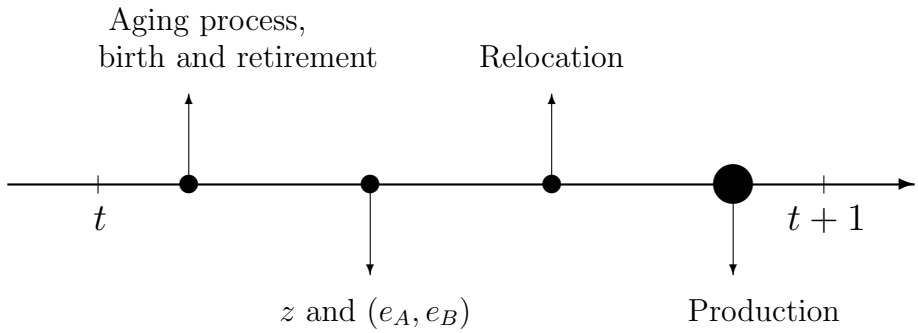
Similarly, if the person works in $B$, her wage will be

$$w_B(z_A, z_B, e_A, e_B, \mu) = \mu + z_B + e_B.$$

Notice that these wage equations do not match with Eq. 1. First, I do not include $G(a)$ in my numerical analysis as it is assumed that the age-earnings profiles are exogenous. Second, I omit the location-specific permanent effect, $\eta_j$, as I compare movers and stayers within a given location.[10]

[10] I abstract from any within-location search process. Therefore, cyclical variation in transitory earnings shocks, $e_j$, may reflect other important shifts such as occupation-, industry- or firm-specific shocks.

Figure 3: Timing of the Events




Each period consists of four stages as shown in Figure 3. At the beginning of each period, some of the young workers become old. At the same time, some of the old workers are retired while young workers are born into the economy. In the second stage, individuals observe their idiosyncratic productivity shocks $(e_A, e_B)$ along with the island-specific technology shocks, $(z_A, z_B)$. In the third stage, after observing these shocks, individuals decide which island to work on for the current period. In the last stage, production takes place. In terms of the duration, only the last stage has a nonzero length.

### 3.2.3 Old workers

For an old worker who worked in the previous period on island $A$, the flow utility of staying in her current location is given by

$$S_A^o(z_A, z_B, e_A, e_B, s, \mu) = \mu + z_A + e_A + \beta(1-\delta^o)EU_A^o(z'_A, z'_B, e'_A, e'_B, s, \mu) \tag{4}$$

where $\beta$ is the time-discount factor. The factor $\beta(1-\delta^o)$ takes into account that in each period, old agents leave the labor market with probability $\delta^o$. If the old person moves to island $B$ and work there for the current period, her lifetime utility is

$$\begin{aligned} M_A^o(z_A, z_B, e_A, e_B, s, \mu) &= \mu + z_B + e_B - T^o(s, \mu) + \\ &\quad +\beta(1-\delta^o)EU_B^o(z'_A, z'_B, e'_A, e'_B, s, \mu). \end{aligned} \tag{5}$$

The expected lifetime utility of a worker who decides whether or not to move is given by

$$U_A^o(z_A, z_B, e_A, e_B, s, \mu) = \max\{S_A^o(z_A, z_B, e_A, e_B, s, \mu), M_A^o(z_A, z_B, e_A, e_B, s, \mu)\}. \tag{6}$$

### 3.2.4 Young workers

Similarly, given the aging probability $\delta^y$, for a young person who worked in the previous period in island $A$, the flow utility of staying in her current location is given by

$$\begin{aligned} S_A^y(z_A, z_B, e_A, e_B, s, \mu) &= \mu + z_A + e_A + \\ &+\beta(1-\delta^y)EU_A^y(z'_A, z'_B, e'_A, e'_B, s, \mu) + \\ &+\beta\delta^y EU_A^o(z'_A, z'_B, e'_A, e'_B, s, \mu) \end{aligned} \tag{7}$$

The expected value of moving to island $B$ is given by

$$\begin{aligned} M_A^y(z_A, z_B, e_A, e_B, s, \mu) &= \mu + z_B + e_B - T^y(s, \mu) + \\ &+\beta(1-\delta^y)EU_B^y(z'_A, z'_B, e'_A, e'_B, s, \mu) + \\ &+\beta\delta^y EU_B^o(z'_A, z'_B, e'_A, e'_B, s, \mu) \end{aligned} \tag{8}$$

Therefore, the maximized expected lifetime utility of a young person will be

$$U_A^y(z_A, z_B, e_A, e_B, s, \mu) = \max\{S_A^y(z_A, z_B, e_A, e_B, s, \mu), M_A^y(z_A, z_B, e_A, e_B, s, \mu)\}. \tag{9}$$

# 4 Estimation

The necessity to take into account the life-cycle effect, the individual-specific fixed effects as well as the self-selection along the transitory income shocks results in a highly intensive computational task. Specifically, solving the model amounts to solving for a stochastic dynamic problem with six state variables when implemented. For computational tractability, I structurally estimate as few parameters as possible, choosing to directly measure and pre-specify the rest. For this purpose, the set of structural parameters, $\Theta$, is divided into two groups: pre-specified, $\Theta_{pre}$, and estimated, $\Theta_{est}$.

## 4.1 Pre-specified parameters

The time period chosen for the numerical simulation is one year. I fix $\beta = 1/1.06$ which reflects the annual interest rate of six percent. Given the mean durations of an individual being young and old, the aging probabilities are given by $\delta^y = 1/15$ and $\delta^o = 1/30$,

respectively. Since $\epsilon$ is the observed transitory shock,

$$\epsilon = \begin{cases} e_A & \text{if the individual works in } A \\ e_B & \text{if the individual works in } B, \end{cases}$$

Due to the self-selection, the observed transitory earnings will be less dispersed than the actual transitory shocks,

$$\text{std}(\epsilon) < \sigma_e,$$

where $\sigma_e$ denotes the standard deviation of the transitory shock, $e_i$ with $i \in \{A, B\}$. Using the March CPS 1981-2005, I find that $\text{std}(\mu + \epsilon)$ is 0.406 and 0.424 for high school- and college-educated groups, respectively. I set $\sigma_\mu^2 = \alpha \text{var}(\mu + \epsilon)$ for some $\alpha > 0$. So, $\alpha$ captures the extent that the permanent earnings component is responsible for the variance of earnings within the same age and schooling level. For example, $\alpha = 1/2$ would indicate that one-half of the within-group earnings variance is due to the transient earnings component and the other half is due to the permanent earnings. To find $\alpha$, I follow Moffitt and Gottschalk (1994, 2002). However, instead of log wage, which they use, I proceed with the rank statistic defined in Section 3.2. Let $\bar{r}_i$ be the mean of $r_{it}$ values of the rank for individual $i$:

$$\bar{r}_i = \frac{1}{T} \sum_t r_{it}.$$

Then, the relative volatility of the two components is given by

$$k = \frac{\text{std}_i(\bar{r}_i)}{\text{mean}_i(\text{std}_t(r_{it} - \bar{r}_i))}.$$

Using the annual hourly wage rate of the heads of the households in PSID[11] from 1968 through 1997, I find that $k = 1.377$ for male high school educated workers and $k = 1.650$ for college educated workers. Since $\alpha = \frac{k^2}{1+k^2}$, $\sigma_\mu = 0.328$ for high school educated workers and $\sigma_\mu = 0.363$ for college educated workers. Further, the variation of the transitory earnings is given by $\text{std}(\epsilon) = \sqrt{(1-\alpha)}\text{std}(\mu + \epsilon)$. Therefore, $\text{std}(\epsilon)$ is equal to 0.239 and 0.220 for high school- and college-educated workers, respectively.

I estimate the persistence of the correlation of individual-specific shocks, $\rho_e$, using

[11] For hourly wage and hours worked, I use the same sample selection criteria as I used in Section 3.2. I also exclude the Survey of Economic Opportunity (SEO) sample which oversamples poor households.

Table 3: Pre-specified Parameters

| Parameter | Value | Description |
|---|---|---|
| Both Schooling Levels | | |
| $\beta$ | 1/1.06 | time discount factor |
| $\rho$ | 0.815 | persistence, local shock |
| $\delta_y$ | 1/15 | probability of aging |
| $\delta_0$ | 1/30 | probability of retirement |
| High School | | |
| $\rho$ | 0.352 | individual productivity (persistence) |
| $\sigma_\mu$ | 0.190 | individual productivity (transitory shock) |
| $\sigma_z$ | 0.051 | regional shock |
| College | | |
| $\rho$ | 0.454 | individual productivity (persistence) |
| $\sigma_\mu$ | 0.182 | individual productivity (transitory shock) |
| $\sigma_z$ | 0.037 | regional shock |

*The normalization details are given in Section 3.4.

PSID and its Geocode Data[12] by restricting our sample to stayers in order to control for regional effects. Clearly, movers are likely to be those who earned less in the previous period at their previous locations. A simple Monte-Carlo exercise shows that, under the normality assumption, the upward bias for the persistence parameter resulting from dropping individuals from the lower end of the within region wage distribution turns out to be negligible due to the low mobility rate. Using the PSID, for the persistence parameter I obtain 0.352 and 0.454 at annual frequency for the high school and college groups, respectively.

To show that using the rank statistic gives a reasonable estimate of the underlying income processes, I consider the following two exercises. First, I pool both the high school and college samples and obtain $k = 1.4307$ and $\rho_e = 0.3678$. These two numbers imply that the extent to which the transitory component is responsible for earnings inequality

[12]Some of the data used in this analysis are derived from Sensitive Data Files of the Panel Study of Income Dynamics, under special contractual arrangements designed to protect the anonymity of respondents. These data are not available from the author. Persons interested in obtaining PSID Sensitive Data Files should contact through Internet at PSIDHelp@isr.umich.edu.

is equal to

$$\frac{\text{var}(\epsilon)}{\text{var}(\mu+\epsilon)} = \frac{1}{k^2+1} = \frac{1}{1.4307^2+1} = 0.328.$$

This is highly consistent with Moffitt and Gottschalk's estimate of one third. Second, the overall persistence from the pooled data is

$$\rho_{US} = \frac{k^2+\rho_e}{k^2+1} = \frac{1.4307^2+0.3678}{1.4307^2+1} = 0.793,$$

which is also remarkably close to Chang and Kim's estimate of 0.809. These consistencies provide additional justification for using a rank statistic in a multi-sectoral setting.

Given $\rho_e$, $\sigma_\nu$ is obtained from the stationarity condition: $\sigma_e^2 = \rho_e^2\sigma_e^2 + \sigma_\nu^2$. I set $\rho_z$ to 0.815. This yearly persistence corresponds to 0.95 in the quarterly frequency. I estimate the magnitude of the regional shock, $\sigma_z$, from Censuses 1980-2000 using mean income as percentage deviation from regional means.[13] The relative magnitude of the local shock is slightly higher for high school educated workers: 0.051 for high school educated workers as opposed to 0.037 for college educated workers. Table 3 summarizes the baseline values of pre-specified parameters.

## 4.2 Estimation method

For the moving cost, I assume a linear function:

$$T^c(s,\mu) = t_{1s}^c + t_{2s}^c\mu$$

for all $s \in \{1,2\}$ and $c \in \{y,\ o\}$. Although the function $T^c(s,\mu)$ is given in a simple linear form, it is flexible enough to reflect the fact that mobility rate differs considerably across age and education levels. This parametric specification along with the above parametrization leaves me with six parameters to estimate for each schooling level:

1. $\rho_x$, the spatial correlation of idiosyncratic shocks;
2. $\sigma_e$, the volatility of the transitory earnings; and
3. $\{t_1^y, t_2^y, t_1^o, t_2^o\}$, the intercepts and slopes of the moving costs.

[13] For geographical differences in overall labor productivity across the U.S., see Bauer and Lee (2005).

for each education level. I estimate these structural parameters by targeting certain moments in the data. Specifically, the estimation entails finding the vector of structural parameters, $\Theta_{est} = \{\rho_x, \sigma_e, t_1^y, t_2^y, t_1^o, t_2^o\}$, which minimizes the distance between $I$ data moments, $\Psi_d = \{\Psi_d^1, \Psi_d^2, \cdots, \Psi_d^I\}$, and their simulated equivalents, $\Psi_s(\Theta_{est}) = \{\Psi_s^1(\Theta_{est}), \Psi_s^2(\Theta_{est}), \cdots, \Psi_s^I(\Theta_{est})\}$. The estimation technique is described in Appendix A. For the data moments, I use the mobility rates, $m$, and conditional mean ranks of annual hourly wage rate of movers, $\tilde{D}^r(1 < \tau_m < 5)$, for each age and education. In addition, I consider the standard deviation of out-migration, std($out$), and the standard deviation of the transitory earnings component, std($\epsilon$), for each education group.

The necessity to measure the wage gap between movers and stayers over a large number of periods requires us to simulate the economy at the individual level rather than keeping track of the evolution of the measures of heterogeneous agents. To simulate the model, I combine discretization of state variables with value function iteration as described in Appendix B.

## 4.3 Results

The results of the estimation of the model specified in the previous section, which I refer to as the benchmark model, are presented in Tables 4 and 5. Under the benchmark model, estimates of the moving costs[14] are approximately 11-16% of annual income depending on age and education. This is much lower than those found in the regional mobility literature.[15] Despite the fact that the mobility rate is higher for old college educated individuals than old high school educated workers, the relative moving cost for old college educated workers is highest among the four groups. For each age and schooling group, mobility decreases with individuals' unobserved ability levels, as is seen from the nonnegative slopes of unobserved ability level, $\mu$.

The most striking difference between the two education levels is the cross-correlation of location-specific shocks: $\rho_x = 0.988$ for high school educated workers and $\rho_x = 0.931$

[14] In Table 4, the moving costs are reported as the percentage of mean annual income for each age and schooling group. Let $T_p$ denote the moving cost expressed in percentage of mean annual income, $\bar{w}$, of the group the individual belongs to. Then, the actual moving cost expressed in dollars is given by $T_d = \bar{w}T_p$.

[15] Kennan and Walker (2006), Lee and Wolpin (2006)

Table 4: Structural Parameters

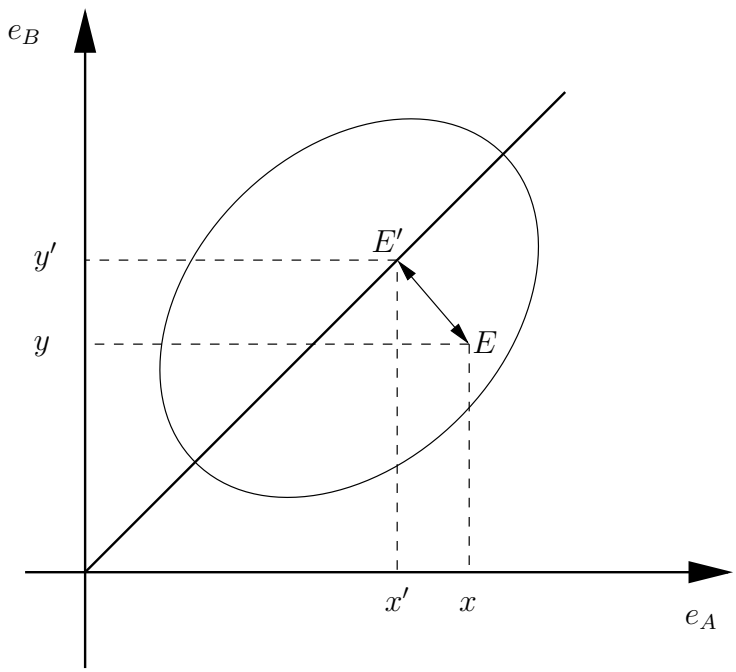
| | Benchmark | | R-1 | | R-2 | | R-3 | |
|---|---|---|---|---|---|---|---|---|
| Restrictions | | | $\rho = 0$<br>$T(\mu) = const$ | | $\rho = 0$ | | $T(\mu) = const$ | |
| Targeted Moments | $m$<br>std($\epsilon$)<br>$\tilde{D}^r$<br>std($out$) | | $m$<br>std($\epsilon$) | | $m$<br>std($\epsilon$)<br>$\tilde{D}^r$<br>std($out$) | | $m$<br>std($\epsilon$)<br>$\tilde{D}^r$<br>std($out$) | |
| | *HS* | *Col* | *HS* | *Col* | *HS* | *Col* | *HS* | *Col* |
| $\rho_x$ | 0.988<br>(0.001) | 0.931<br>(0.002) | | | | | 0.985<br>(0.002) | 0.929<br>(0.002) |
| $\sigma_e$ | 0.242<br>(0.014) | 0.223<br>(0.020) | 0.248<br>(0.008) | 0.231<br>(0.011) | 0.246<br>(0.007) | 0.227<br>(0.010) | 0.244<br>(0.010) | 0.226<br>(0.013) |
| $t_1^y$ | 0.108<br>(0.002) | 0.119<br>(0.001) | 0.731<br>(0.0006) | 0.564<br>(0.000) | 0.656<br>(0.002) | 0.587<br>(0.0004) | 0.121<br>(0.002) | 0.129<br>(0.001) |
| $t_2^y$ | 0.018<br>(0.000) | 0.005<br>(0.001) | | | 0.086<br>(0.0003) | -0.016<br>(0.0006) | | |
| $t_1^o$ | 0.124<br>(0.001) | 0.137<br>(0.002) | 0.834<br>(0.001) | 0.718<br>(0.0003) | 0.851<br>(0.001) | 0.786<br>(0.0004) | 0.129<br>(0.001) | 0.145<br>(0.002) |
| $t_2^o$ | 0.023<br>(0.004) | 0.025<br>(0.006) | | | 0.040<br>(0.002) | 0.048<br>(0.005) | | |

for college educated workers. To measure the effect of the differences between the two values on the cross-sectional volatility of the transitory earnings, let us consider Figure 4. $E$ denotes an arbitrary pair of transitory shocks, $(x, y)$. Its projection on the indifference curve is depicted by $E'$. Let $(x', y')$ be the shocks associated with $E'$. The cross-sectional volatility of the transitory earnings is measured by the variance of the distance between $E$ and $E'$. Since $x' = y' = \frac{x+y}{2}$, the distance between the two points is $|EE'| = \frac{|x-y|}{\sqrt{2}}$. Therefore, the relative variation of the location-specific component is given by

$$\frac{\text{std}(|EE'|)}{\sigma_e} = \frac{\sqrt{\frac{\text{var}((\rho_x - 1)e_A + u_A)}{2}}}{\sigma_e} = \sqrt{1 - \rho_x}.$$

The estimates imply that approximately 10% of the transitory earnings variation among high school educated individuals is due to the pure location-specific effect. However, for college graduates, the location-specific effect amounts to 25% of the transitory earnings variation. More importantly, the cross-sectional variation of the transitory earnings shocks

Figure 4: Cross-Sectional Volatility of Transitory Shocks




$E$ denotes an arbitrary pair of transitory shocks, $(x, y)$. Its projection on the 45° line is $E'$. The variation of the location-match effect is given by the variation of $|EE'|$.

of college educated workers is more than twice that of high school educated workers:

$$\frac{\text{std}(|EE'|)_{COL}}{\text{std}(|EE'|)_{HS}} = \sqrt{\frac{1-\rho_{x2}}{1-\rho_{x1}}}\frac{\sigma_{e2}}{\sigma_{e1}} \simeq 2.2.$$

These exercises show that the strength of the cross-correlation coefficients is the main factor responsible for the large observed differences in mobility patterns between the two education levels. As is seen from Table 1, the gross mobility of college educated workers is twice as high as that of high-school graduates. However, net mobility is much higher for high-school educated workers. These differences indicate that high-school educated workers are more susceptible to local labor market disturbances than college educated workers. In order to obtain highly susceptible net mobility, the relative effect of idiosyncratic labor shocks must be much lower and hence the cross-correlation for idiosyncratic labor shocks must be very high in the model. This indicates that labor market thickness decreases

Table 5: Simulated Moments

| | Benchmark | | R-1 | | R-2 | | R-3 | | Data | |
|---|---|---|---|---|---|---|---|---|---|---|
| | HS | C | HS | C | HS | C | HS | C | HS | C |
| std($\epsilon$) | 0.241 | 0.219 | 0.238 | 0.220 | 0.237 | 0.219 | 0.238 | 0.218 | 0.239 | 0.220 |
| $m_y$ | 0.082 | 0.173 | 0.087 | 0.172 | 0.109 | 0.154 | 0.086 | 0.168 | 0.089 | 0.171 |
| $m_o$ | 0.046 | 0.115 | 0.035 | 0.075 | 0.039 | 0.041 | 0.040 | 0.092 | 0.036 | 0.072 |
| std($out$) | 0.021 | 0.015 | 0.005 | 0.007 | 0.010 | 0.010 | 0.016 | 0.013 | 0.021 | 0.015 |
| $\tilde{D}_y^r$ | 0.430 | 0.514 | 0.563 | 0.527 | 0.472 | 0.546 | 0.503 | 0.532 | 0.434 | 0.514 |
| $\tilde{D}_o^r$ | 0.481 | 0.550 | 0.545 | 0.564 | 0.473 | 0.563 | 0.521 | 0.578 | 0.478 | 0.560 |

with education: for high school graduates, employment opportunities do not vary much across different markets whereas for college educated individuals, different markets offer different opportunities.

As we discussed earlier, the high school educated movers earn less than their non-migrant counterparts. Additionally, newly-arrived college educated workers earn more compared with the college-educated non-movers. The estimation results offer the following explanation for these differences across education levels. For high school educated workers, the earnings gap of movers is smaller due to the lower moving costs of lower ability workers. However, for college educated workers, this effect of permanent ability level is offset by the positive sorting along the transitory shock.

To measure the biases caused by the assumption that location-specific shocks are not cross-correlated, I conduct two counterfactual experiments while imposing certain parametric and moment restrictions to the benchmark model. First, I estimate the model while restricting the cross-correlation of location-specific shocks to zero, targeting only gross mobility and variation of the transitory earnings component. Second, I restrict the cross-correlation of location-specific shocks to zero, targeting all the moments used in the benchmark specification. In Tables 4 and 5, the two restricted models are denoted by R-1 and R-2, respectively.

Under the first restriction, the movings costs are approximately 60-80% of annual income. These estimates are much larger than what is obtained using the benchmark specification. Although the model performs well along the targeted moments, it has counterfactual implications on the earnings gaps between movers and stayers: For both young and old high school groups, newly arrived individuals earn much higher than the local residents. Also, net mobility is much lower than its empirical counterparts. Under the second restriction, moving costs still remain large. However, due to the inclusion of net mobility into the set of the targeted moments, the estimation gives slightly lower moving costs compared with R-1.

These experiments indicate that more flexible moving costs help improve on earnings differences between movers and stayers. To gain further insights into the effects of the slope of the moving costs, I conduct one more experiment while restricting the moving costs to be orthogonal to individuals' permanent ability levels. This restriction is denoted by R-3 in Tables 4 and 5. As seen from Table 5, the main direction along which the models with the orthogonality restriction $T(\mu) = const$, i.e R-1 and R-3, do not perform well when compared to the benchmark model is the earnings gap. Contrary to the empirical facts discussed earlier, the movers in these two restricted models earn more than the stayers across age and education groups. The reason is that individuals from different permanent ability levels are equally mobile under this restriction and therefore, the positive sorting along the transitory shocks raises movers' wage relative to local residents.

# 5 Conclusion

This paper was motivated by the realization that the migrants are not randomly selected from the population and that migration decisions are reversible. It documents new evidences on patterns of earning differences between movers and stayers for age and schooling groups. The paper explains these differences using a dynamic Roy model in a directed search setting. The model extends the earlier works of self-selection and internal migration along two main dimensions: a) explicit treatment of both gross and net mobility, and b) a richer structure on the transitory earnings shocks across locations. The results of

this paper show that assuming uncorrelated transitory shocks in modeling sectoral mobility may lead to serious biases in the structural parameters. The model makes a strong prediction that market thickness decreases with education: for college educated workers different markets offer different opportunities whereas for less educated workers opportunities do not much across local markets. Despite it is simplicity, the model establishes a link between the two important literatures: regional mobility and individual labor income processes.

It would be of interest to repeat the exercises in this paper for occupational or inter-industry mobility. This may shed light on the extent to which individuals' labor income is influenced by occupation- and industry-specific skills. Another interesting exercise would be to estimate the model using European data and compare the results with those obtained in this paper. This may help us understand the differences in labor market outcomes between the two economies. For instance, it can help us determine whether higher labor mobility in the U.S. is due to its lower moving costs or higher earnings volatility.

In the current model, the moving costs reflect both punitive and non-punitive costs associated with switching locations. The prediction that high-ability workers have higher moving costs within each age and schooling levels could be linked to home ownership, housing quality and spousal attachment to the local market. Future research should look at these important variables. Due to the data limitations, I conducted my analysis at the level of the U.S. Census-Divisions abstracting from other smaller sub-national geographic characteristics. Nevertheless, my study can be viewed as an important step towards understanding how cross-sectional earnings variance and sectoral mobility are related.

# References

[1] Altonji, Joseph G., and Lewis M. Segal. 1996. "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 14(3): 353-66.

[2] Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants," *American Economic Review,* 77(4): 531-53.

[3] Borjas, George J., Stephen G. Bronars, and Stephen J. Trejo. 1992. "Self-Selection and Internal Migration in the United States." *Journal of Urban Economics*, 32(2): 159-185.

[4] Chang, Yongsung, and Sun-Bin Kim. 2006. "From Individual to Aggregate Labor Supply: A Quantitative Analysis based on a Heterogeneous-Agent Macroeconomy." *International Economic Review*, 47: 1-27.

[5] Dahl, Gordon B. 2002. "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets," *Econometrica*, 70: 2367-2420.

[6] Galindev, Ragchaasuren, and Damba Lkhagvasuren. 2008. "Discretization of Highly Persistent Correlated AR(1) Shocks." Working Papers of Department of Economics, Concordia University, 08012.

[7] Gomme, Paul, and Peter Rupert. 2005. "Per Capita Income Growth and Disparity in the United States, 1929−2003." *Policy Discussion Paper*, Federal Reserve Bank of Cleveland.

[8] Greenwood, Michael J. 1997. "Internal Migration in Developed Countries." In *Handbook of Population and Family Economics,* ed. Mark R. Rosenzweig and Oded Stark, Vol. 1B. New York: North Holland.

[9] Guvenen, Fatih. 2009. "An Empirical Investigation of Labor Income Processes." *Review of Economic Dynamics,* 12(1): 58-79.

[10] Heckman, James J., and Guilherme L. Sedlacek. 1990. "Self-Selection and the Distribution of Hourly Wages," *Journal of Labor Economics*, 8: S329-63.

[11] Heckman, James, Anne Layne-Farrar, and Petra Todd. 1996. "Human Capital Pricing Equations with an Application to Estimating the Effect of Schooling Quality on Earnings." *The Review of Economics and Statistics,* 78: 562-610.

[12] Kennan, John, and James R. Walker. 2008. "The Effect of Expected Income on Individual Migration Decisions." Working Paper, University of Wisconsin-Madison.

[13] Lee, Donghoon, and Kenneth I. Wolpin. 2006. "Intersectoral Labor Mobility and the Growth of the Service Sector," *Econometrica,* 74: 1-46.

[14] Lkhagvasuren, Damba. 2007. "Local Labor Market Dynamics with Net and Gross Mobility: Implications on Unemployment and Wages." Ph.D. Dissertation, University of Rochester.

[15] Lkhagvasuren, Damba. 2009. "Large Locational Differences in Unemployment Despite High Labor Mobility: Implications on Aggregate Unemployment." Working Paper, Concordia University.

[16] McLaughlin, Kenneth J., and Mark Bils. 2001. "Interindustry Mobility and the Cyclical Upgrading of Labor," *Journal of Labor Economics,* 19: 94-135.

[17] Moffitt, Robert A., and Peter Gottschalk. 1994. "The Growth of Earnings Instability in the U.S. Labor Market." *Brookings Papers on Economic Activity,* 1994(2): 217-272.

[18] Moffitt, Robert A., and Peter Gottschalk. 2002. "Trends In The Transitory Variance Of Earnings In The United States." *The Economic Journal*, 112: C68-C73.

[19] Nakosteen, Robert A., and Michael A. Zimmer. (1980). "Migration and Income: The Question of Self Selection," *Southern Economic Journal*, 46(3): 840-851.

[20] Roy, A. D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers,* 3: 135-146.

[21] Tauchen, George. 1986. "Finite State Markov-Chain Approximation to Univariate and Vector Autoregression." *Economics Letters,* 20: 177-181.

[22] Wozniak, Abigail. 2006. "Differences in the Migration Responses of Young Workers to Local Labor Market Conditions." *IZA Discussion Paper Series,* DPNo. 1954.

# Appendixes

## A. Details of the estimation method

This appendix describes the SMM estimation of $\Theta_{est}$. Our parameter estimates $\hat{\Theta}_{est}$ are obtained by minimizing the weighted distance,

$$L(\Theta_{est}) = (\Psi_s(\Theta_{est}) - \Psi_d)W(\Psi_s(\Theta_{est}) - \Psi_d)', \tag{10}$$

where $W$ denotes the weighting matrix. Using a Monte-Carlo study, Altonji and Segal (1996) show that the optimal weighting matrix is often dominated by equally weighted matrix when the sample size is small. Since there are nine Census Divisions and the key empirical moments are measured from Census 1980, 1990 and 2000, there are only $9 \times 3 = 27$ observations at the regional level to measure $std(out)$. Therefore, following their recommendation, I use an identity matrix for $W$.[16] The estimator $\Theta_{est}$ is consistent and asymptotically normal with the asymptotic covariance matrix

$$V = (Q'Q)^{-1}Q' \frac{\Omega}{\tau} Q(Q'Q)^{-1},$$

where $Q = \frac{\partial \Psi_s(\Theta_{est})}{\partial \Theta'_{est}}$ and $\Omega = n\text{var}(\Psi_s(\Theta_{est}) - \Psi_d)$. $\hat{\Omega}$ is estimated as a time average of the simulated moments

$$\hat{\Omega} = \frac{1}{\tau - 1} \sum_{t=1}^{\tau} (\Psi_{st}(\hat{\Theta}_{est}) - \Psi_d)(\Psi_{st}(\hat{\Theta}_{est}) - \Psi_d)'.$$

## B. Numerical method

To solve for the policy function governing the individual migration decision, I use the value function iteration method. With the extensive discretization of various shocks, the value function iteration method reduces to simple manipulation of matrices. I approximate the stochastic process for the local technology shock, $z$, by a 15-point Markov chain. To construct the transition matrices associated with the single AR(1) shocks, I follow Tauchen (1986). For the two cross-correlated $AR(1)$ shocks, I use $50 \times 50$ grid points. The transition matrix of the cross-correlated shocks is constructed using the method proposed by Galindev and Lkhagvasuren (2008). The individual-specific effect is accounted

[16] Guvenen (2009) also uses the same method to study labor income processes in PSID.

by two levels of $\mu$: $\mu_L = -\sigma_\mu$ and $\mu_H = \sigma_\mu$. The necessity to measure the relative performance of movers for four consecutive periods requires us to simulate the economy at the individual level rather than keeping track of the evolution of measures of observationally identical agents. Thus, the simulation of the economy is computationally highly intensive. I simulate the economy of 5000 agents for 500 periods. When I calculate the simulated moments I discard the first 150 periods.

I solve the minimization problem (10) by the Nelder-Mead simplex method. For the initial guess of the minimization, I use simulated annealing. Simulated annealing is a probabilistic algorithm for the global optimization problem in a high-dimensional search space. Each step of the algorithm replaces the current solution with a random solution in its vicinity with a probability that depends on the difference between the two function values, as well as a parameter $K$ that is gradually tuned. When $K$ is large the current solution changes almost randomly. However, it goes increasingly "downhill" as T goes to zero. Allowing the algorithm to make "uphill" moves saves the method from becoming trapped at local minima.