Understanding Gross Worker Flows Across U.S. States

Daniele Coen-Pirani*
Carnegie Mellon University

July 14, 2009

Abstract

A dynamic general equilibrium model of migration is developed to explain the main features of geographic mobility of workers across U.S. states. Models of net flows only cannot explain the positive cross-sectional correlation between gross inflow and outflow rates. The dynamics of migration flows is driven by local productivity shocks and idiosyncratic shocks to the match between a worker and a location. The latter is revealed only after migration has occurred. Thus, migrating workers have a higher propensity to migrate again than incumbent workers and locations that attract high numbers of migrants also tend to experience high outflow rates.

Keywords: Migration, States, Employment, Gross Flows, Net Flows, Wages, Rents, Productivity, Land.

JEL Classification: J00, J60, R13.

*Address: Tepper School of Business, 5000 Forbes Avenue, Pittsburgh, PA 15213. E-mail: coenp@andrew.cmu.edu. Phone: (412) 268-6143. Fax: (412) 268-7064. Thanks, without implication, to seminar participants at Carnegie Mellon, UT Austin, SUNY Albany, Cleveland Fed, Kansas Fed, UQAM, LABORatorio Revelli, the 2006 Midwest Economics Association Meetings, the 2006 Society for Economic Dynamics Meetings, the 2007 Texas Monetary Conference, the 2007 Midwest Macro Meetings, the 2007 Econometrics Society North American Summer Meetings, and to Nathan Balke, Yoonsoo Lee, Chris Flinn, Holger Sieg, Bruno Contini, Fallaw Sowell, Marla Ripoll, Roc Armenter, Maria Ferreyra, Dennis Epple, Alexis Leon, Iourii Manovski, Rui Castro, B. Ravikumar, Jordan Rappaport, Rodi Manuelli, Ted Temzelides for useful discussions and suggestions. Richard Lowery provided able research assistance.

1 Introduction

Migration of workers across states is the main factor behind the large and persistent dispersion in the growth rates of employment across U.S. states (Blanchard and Katz (1992)). Yet, the process by which population is reallocated among geographic areas within a country is not well understood yet. This paper argues that in order to improve this understanding it is important to consider *both* gross and net flows of workers. First, gross flows are large relative to net flows. For example, between 1995 and 2000 the average state gained or lost about 2.1 percent of its 1995 population. In the same period, the average state experienced a combined inflow and outflow of population of about 17 percent of its 1995 population. Second, the separate consideration of each gross flow allows us to sharpen our understanding of the mechanism by which a location gains employment relative to other states. Is this due to higher-than-average gross inflows or, alternatively, to lower-than-average gross outflows of workers? Third, models of net flows (e.g. Lucas and Prescott (1974)) cannot account for the fact that locations that attract large numbers of internal migrants relative to their population also tend to experience large out-migration in relative terms. For example, Figure 1 plots "residual"[1] gross outflow rates of workers from U.S. states against gross inflow rates using pooled data from the 1970-2000 U.S. Censuses. The correlation coefficient between these two variables is 0.73 and statistically significant. A similar relationship holds also across metropolitan areas. For example, the cross-sectional correlation between inflow and outflow rates of workers across the 100 largest U.S. metropolitan areas in 2000 is 0.41 and also statistically significant.

This paper has two goals. The first one is characterize the main cross-sectional and time-series properties of worker flows, state-level wages, and rents using data from the U.S.

[1] Worker flows are constructed using information on the worker's state (or metropolitan area) of residence at the time of the Census and her state (or metropolitan area) of residence 5 years before the Census year. Section 2 and Appendix A describe the data and provide details on how inflow and outflow rates are computed. The qualifier "residual" refers to the fact that both inflow and outflow rates have been adjusted for composition effects and for the potential influence of gravity forces (distance and size) on bilateral flows. The variables plotted in Figures 1 (as well as in Figures 2 and 3) are residuals from these regressions. From now on I will refer to adjusted worker flows simply as worker flows.

Census. The second goal is to account for these stylized facts using a dynamic general equilibrium model of worker migration.

Census data reveal interesting information about the features of migration flows and the nature of the shocks responsible for employment reallocation at the state-level. First, there is significantly less variation in gross inflow than in gross outflow rates across U.S. states. States that experience relative losses of population do so mainly because of low inflows rather than because of especially high outflows. Second, the micro data from the Census reveal that workers moving into a U.S. state tend to be observationally similar to workers migrating out of it. In particular, these workers tend to have a similar age and education and to work in similar industries. This evidence runs counter to the argument advanced by Larry Sjaastad (1962, 1961), who in his classic work on migration stated that the "somewhat paradoxical relation between gross in- and out-migration may be substantially an aggregation problem," meaning that "gross migration reflects the degree to which the labor force is being reshaped by changing demand and supply conditions among industries." Third, net and gross flows of workers are very persistent across different Census years. Thus, while state-level business cycles might give rise to some of the observed migration flows, much of the observed reallocation of employment across states is due to shocks characterized by much lower frequency than cyclical shocks. Fourth, in the cross-section, states with relatively high earnings are also characterized by relatively high land prices. This pattern points to the importance of labor demand rather than labor supply shocks in driving geographic reallocation of workers (Roback (1982)).

I interpret these stylized facts using an equilibrium model of gross and net migration flows. The model economy is composed by a set of local labor and land markets that are hit by shocks to total factor productivity. Ex-ante identical workers are assumed to be mobile across locations while firms use constant-returns-to scale technologies in labor and land to produce an homogeneous good. At a point in time, a location typically experiences both gross inflows and gross outflows of workers because of a matching effect: workers can observe aspects of

their match with a location only after having migrated there. The calibrated version of the model accounts relatively well for the cross-sectional and time-series properties of net and gross worker flows and for the relationship between worker flows and the local prices of labor and land.

In particular, the model reproduces the positive relationship between inflow and outflow rates observed in the data through the selection effect associated with ex-post heterogeneity. Locations hit by positive productivity shocks attract a relatively large number of internal migrants as a fraction of their population. Workers who migrate in one period have a higher propensity to migrate again in the following period than incumbent workers who do not migrate. Hence, locations characterized by high inflow rates in the current period experience relatively large gross outflow rates in *future* periods. The *contemporaneous* correlation between inflow and outflow rates obtains because gross inflow rates are persistent over time. I test this implication of the model using data on gross worker flows across several Census decades. The results are consistent with the model's mechanism: in a regression of current outflow rates on both contemporaneous and lagged inflow rates, the estimated coefficient on lagged inflows is positive and significant, while the one on current inflows is negative.

This paper is related to several literatures.[2] In macroeconomics, the general equilibrium approach to "migration", broadly meant to include moves from one employment status or sector to another, was pioneered by Lucas and Prescott (1974)'s island model of the labor market. Jovanovic and Moffitt (1990) consider an overlapping-generations version of the Lucas-Prescott model to study gross flows of workers across industries. At the micro-level,

[2] There is other related work in the macro and regional literatures that should be mentioned. Kambourov and Manovskii (2009) use the Lucas-Prescott model to study trends in occupational mobility and wage inequality. Alvarez and Veracierto (2000 and 2006) analyze labor market policies. Alvarez and Shimer (2007) study different types of unemployment. Topel (1986) extend the Lucas-Prescott model to allow for heterogeneity in experience among workers. Van Nieuwerburgh and Weill (2006) introduce housing. Differently from these contributions based on a competitive setting, Lkhagvasuren (2005) adopts a Mortensen-Pissarides style of model to explain the existence of persistent differentials in unemployment rates across U.S. states. Morris and Ortalo-Magne (2007) study the joint distribution of housing prices and wages across metropolitan areas. The paper is also related to the research on the determinants of population flows within the U.S. developed by Greenwood and Hunt (1984), Treyz et al. (1993), Rappaport (2004), and Armenter and Ortega (2007).

Kennan and Walker (2006) estimate a partial equilibrium model of optimal migration across U.S. states. In their model, as in my model and in Jovanovic and Moffitt (1990), workers need to migrate to a location in order to observe the value of their idiosyncratic match with it. One of Kennan and Walker's contributions is to study return migration, an important issue from which I abstract in this paper in order to study migration decisions in general equilibrium. In my model workers care about consumption of goods and consumption of an immobile factor, land. From this perspective, the model generalizes Roback (1982)'s static model of the effect of local amenities on local land prices and wages to a dynamic setting. The presence of land in the utility function allows the model to explain why net migration flows are not directed toward locations where the price of labor is relatively high. On the empirical front, the paper builds on the seminal contribution of Blanchard and Katz (1992), who document the existence of very persistent differences in employment growth rates across U.S. states. Related work by Bartik (1991) and Davis, Loungani, and Mahidhara (1991) studies the effect of labor demand shocks on yearly deviations of local employment from location-specific long-run trends. In doing so these authors use local industry shares to construct measures of local labor demand shocks based on national trends in sectoral employment. My focus on location-specific, as opposed to industry-specific, productivity shocks reflects the low-frequency nature of the Census data and is therefore consistent with the shift-share approach taken by the papers cited above.[3]

The rest of the paper is organized as follows. Section 2 describes the data and the stylized facts. Section 3 presents the model. Section 4 describes the calibration. Section 5 discusses the model's fit. Section 6 performs sensitivity analysis. Section 7 provides additional empirical evidence on the mechanism of the model. Section 8 concludes. The

[3] For example, the decline of the steel industry in Pennsylvania also led to lower demand for local services (education, health, etc.) and construction. While the original shock affected one industry, over a sufficiently long time-horizon the disaggregated data show a broad decline in employment across most sectors of the local economy. Using yearly employment data on one-digit industry employment by U.S. state from the Bureau of Labor Statistics, 1969-2001, I find that *national* trends in employment shares by industry can explain about 40 percent of the variance of employment growth across states over a one year period. However, over a five year period this figure drops to 7 percent, and over a ten year period to only 1 percent.

appendices offer a more detailed description of the data, of the construction of the flow variables, and of the algorithm used to solve and calibrate the model.

2 Data and Stylized Facts

In this section I briefly describe the data used in the paper and then present the main stylized facts.

2.1 The Data

I use the IPUMS data extracts from the U.S. Census of Population for 1970, 1980, 1990, and 2000 (Ruggles et al. (2004)).[4] In order to focus on geographic mobility that is not motivated by college attendance or retirement, I restrict attention to individuals who are between 28 and 60 years of age and in the labor force at the time of the Census. The sample includes both U.S.-born workers as well as foreign-born ones who immigrated to the U.S. at least five years prior to the Census year. This restriction guarantees that aggregate net flows of workers equal zero.[5] The Census provides information on an individual's state of residence at the time of the interview and five years before. I use this mobility variable to construct state-level measures of inflows and outflows of workers. Prior to doing so, and in order to control for composition effects, I regress an individual's mobility choice of staying vs. moving on a set of demographic controls in each Census year. The residuals from this regression are used to construct bilateral gross inflow and outflow rates of population for the 48 contiguous U. S. states. Rates are computed by dividing the bilateral flows by the number of workers who are surveyed in the Census year and report living in that state five years before. In order to control for the effect of gravity forces such as distance and size on population flows, the bilateral gross flow rates are regressed in each Census year on measures

[4] The data is available online at `http://usa.ipums.org/usa/`. Appendix A contains detailed information on the data, issues of sample selection, and on the construction of the variables used in the paper.

[5] In Appendix A.6, I argue that recent immigration from outside of the U.S. does not significantly affect the stylized facts presented in Section 2.2.

of distance between states, states' sizes, and the presence of a border. The residual bilateral gross flow rates are then aggregated to create a gross inflow rate and a gross outflow rate for each of the 48 states for each Census year. Net flow rates are defined as the difference between these gross inflow and outflow rates. These rates are used to compute the statistics reported in the following section.

2.2 Stylized Facts

In what follows I introduce the main features of the data on worker flows using simple descriptive statistics. The moments in the tables following Table 1 are constructed by using all state-year observations and weighting each state by its relative population, but results do not depend on the use of weights.

2.2.1 Magnitude of Gross Flows

I start by summarizing the magnitude of gross and net flows. The second column of Table 1 reports gross worker reallocation - defined as the sum of gross inflows and outflow rates - across U.S. states. The third column represents the average absolute value of net flow rates across U.S. states, while the fourth column reports the excess reallocation of workers, defined as the difference between columns two and three. Notice that gross flows are large relative to net flows. On average, a state gains or loses about 2.8 percent of its workforce in a five-year period due to migration. In the same period, the average state experiences a combined inflow and outflow of workers of 16.3 percent of its workforce.[6]

[6] Notice that this measure is likely to underestimate the true mobility rate because a subset of workers might move more than once in a five-year period. In fact, according to the 2000 American Community Survey, the one-year migration rate is about 2.2 percent. In order to account for these time-aggregation effects, in the quantitative part of the paper I assume that one period in the model is equal to one year. I then use the data generated by the model to compute the five-year migration rates, just as in the Census.

2.2.2 Cross-Sectional Dispersions and Correlations Among Worker Flows

There are considerable cross-sectional differences among states in terms of net and gross worker flow rates, with gross inflow rates displaying the largest differences, followed by net flow rates, and gross outflow rates (see Table 2). This ranking of variances implies that gross inflow and outflow rates as well as gross and net inflow rates are positively correlated in the cross-section.[7] These correlations are reported in Table 3.

These basic relationships between outflow, inflow, and net flow rates are also evident from the scatter plots in Figures 1 and 2 which represent gross inflow and outflow rates and net and gross inflow rates, respectively. The sign of the relationship between net flow and gross outflow rates cannot be deduced from the ranking of variances in Table 2, instead. It turns out that, as shown in Figure 3, there is a U-shaped relationship between these two variables, with the resulting correlation between them being positive, but close to zero.[8] It follows that states that tend to lose workforce due to migration do so by experiencing lower than average inflows, rather than higher than average outflows. For example, in the 2000 Census North Dakota and Nevada have both very large outflow rates (around 13-14 percent), even if they are at the opposite extremes of the distribution of net flow rates.

The positive cross-sectional correlation between gross inflow and gross outflow rates reported in Table 3 might be symptomatic of a changing industry or demographic mix of states' workforce (Sjaastad (1962)). To answer the question of whether gross worker flows occur mainly within or between demographic and industry groups, I divide the 2000 Census sample into 385 demographic groups defined according to age, education, and industry.[9] I then compute gross and net flow rates for each state and for each demographic group. A way to consider exclusively within-group flows is to calculate the cross-sectional correlations among worker flow rates separately for each of the 385 demographic groups. Table 4 reports

[7] This follows easily from the definition of net flow rates as the difference between gross inflow and outflow rates and from the definition of covariance.

[8] Regressing gross outflow rates on net flow rates and the square of net flow rates yields positive and significant coefficients on both regressors.

[9] See Appendix A.3 for a detailed explanation of how these groups were constructed.

the 25th, 50th, and 75th percentile of the distribution of these correlation coefficients for the 2000 Census. Notice that the correlations in Table 4 display a similar pattern as the ones in Table 3. In particular, the within-group correlations between inflow and outflow rates, while lower than the value reported in Table 3, are still positive and statistically significant for the great majority of demographic groups. Thus, for a given state, incoming workers tend to have a similar age and level of education and to work in the same industry as outgoing ones. This result is consistent with Miller (1967) who finds that inflow and outflow rates across U.S. metropolitan areas are positively correlated even after controlling for workers' sex, race, and occupation.

A complementary exercise which confirms these results is to compute for each U.S. state the fraction of excess worker reallocation that is due to between-group employment shifts. The population-weighted average of this measure across U.S. states for the 2000 Census is 8 percent, suggesting that most flows occur within the demographic/industry groups described above. Similar results obtained for the other Census years.

Last, to further assess the nature of worker flows, it is possible to compare weekly earnings of workers who are migrating into a given state with the weekly earnings of workers who are leaving that same state. If outgoing and incoming workers belong to the same demographic groups we should not observe sizable differences between those. Without controlling for any observable worker characteristic, the average weekly labor earnings of a worker moving out of a state are about three percent lower than the weekly earnings of a worker moving into that same state. This difference is quite small, further confirming the within-group nature of gross migration flows.

2.2.3 Worker Flows Over Time

Thus far I have focused on the cross-sectional correlations among worker flows. The advantage of using several Census years is that it is possible to assess the persistence of worker flows over time. Table 5 reports, for each type of flow, its first order autocorrelation coef-

ficient across Census years, computed by pooling all state-year data points together. The autocorrelation coefficients imply that worker flows are very persistent across decades, suggesting that low-frequency shocks play a major role in reallocating employment across U.S. states. This finding is in accordance with the evidence on states' employment and population growth presented by Blanchard and Katz (1992) and Rappaport (2004).

2.2.4 Worker Flows and State-Specific Wages and Rents

The Census data also provide useful information on individual wages and rents paid. The individual price information can be aggregated at the state level after controlling for observables.[10] Table 6 presents the most salient facts about worker flows, state-level nominal weekly wages, and rents, where the latter variable is meant to capture the price of a local immobile input used for both production and consumption purposes. From the table it is clear that states with relatively higher average wages are also characterized by higher average rents. The correlation between wage and rent levels and net population flows is close to zero, instead, as also documented by Blanchard and Katz (1992). Finally, both differences in average wages and rents across states are quite persistent across Census decades.

3 Model

In order to account for these stylized facts I introduce a general equilibrium model of migration. The model builds on the island-model of the labor market developed by Lucas and Prescott (1974) and on Roback (1982)'s static analysis of worker and firm location decisions. The force that drives the dynamics of the local labor and land markets in the model is a persistent shock to total factor productivity. The latter generates temporary increases in local wages and land prices that are then followed by net inflows of workers. Simultaneously,

[10] State-level labor earnings and land rents are measured as state fixed-effects in hedonic regressions of individual earnings and rents on observable characteristics of workers and renter-occupied housing units. See Appendices A.4 and A.5 for details. The cross-sectional correlations in the data were computed by pooling all state-year observations together and controlling for year dummies.

idiosyncratic match shocks give rise to workers' gross flows. In equilibrium, the value of migrating from one labor market to another is pinned down by the requirement that aggregate net flows of workers are zero.

3.1 Description

Production and Firms. The economy is populated by a continuum of locations ("islands") of measure one. Each location is endowed with one unit of a local and immobile factor of production ("land"). As in Roback (1982), land is used in production and supplies a service to the consumer. There is only one good in this economy whose price is normalized to unity. The good is produced in each location by a large number of competitive firms with the constant-returns-to-scale production technology $zF(y, l^f)$. The production input y represents the units of labor located in the island while l^f is land used in production.[11] The production function F is concave and satisfies the standard Inada conditions. The variable z represents total factor productivity in the island. Firms maximize static profits upon observing the productivity shock z. Profit maximization requires that they equalize the marginal product of labor and land to their local relative prices, denoted by w and r respectively.

Exogenous Shock. The law of motion for z' depends on z and another random variable ε. Letting $\zeta = (z, \varepsilon)$, the exogenous state vector for a location is assumed to evolve over time according to a stationary Markov process with transition function $Q(\zeta', \zeta)$. The transition function $Q(\zeta', \zeta)$ takes the following form:

$$z' = z\varepsilon', \tag{1}$$

$$\varepsilon' = z^{\psi-1}\varepsilon^{\rho}u' \tag{2}$$

[11] The production function could also be made a function of physical capital. Under the assumption of perfect capital mobility, the latter would move to equalize its rate of return across locations. In this case, the function F in the text should be interpreted as the reduced-form production function obtained by solving out for the optimal amount of capital in a location and replacing the latter back into the original production function. I take physical capital into account when calibrating the model in Section 4.

where $\psi < 1$, $\rho < 1$, and u' is independently and identically distributed both over time and across locations according to a lognormal distribution with mean one and variance $(\exp\{\sigma_u^2\} - 1)$. The specification of the exogenous shocks in equations (1)-(2) is somewhat non-standard because it assumes that the growth rate of productivity ε' is persistent. This is necessary in order to generate persistent net flows in a Lucas-Prescott type of model such as this one. If ρ were equal to zero, the autocorrelation of net flow rates would be close to zero, which is strongly at odds with the data (see Table 5).[12] The level of z (with the assumption $\psi < 1$) is introduced in equation (2) in order to keep the process for z stationary.

There are two complementary ways of thinking about the nature of the shock z. According to the first, z captures location-specific productivity differentials that in principle affect different industries in the same way (see Holmes (1998) and Clark (1998)).[13] According to a second view, z is not a productivity shock to all sectors in a location, but instead represents, in a reduced-form way, the effect that sectoral shocks have on the local economy, once local intersectoral interactions are accounted for. Finally, Appendix B shows that a version of the model with constant productivity but stochastic "amenity" shocks (i.e., the introduction of air conditioning) that affect workers' utility in a location would have exactly the same implications for migration flows across locations as the model considered in this section.

Preferences, Idiosyncratic Shocks, and Budget Constraint. The economy is populated by a measure one of infinitely-lived agents. An agent's instantaneous utility function is assumed to take the following quasi-linear form:

$$u = \upsilon \times (c + \phi(l^c)), \tag{3}$$

[12] An alternative way of obtaining persistent net flows would be to introduce less-than-perfect capital mobility in the model (see e.g. Rappaport (2004)), as opposed to the current setting in which capital is assumed to be perfectly mobile. The parameter governing capital adjustment costs would then determine the extent of autocorrelation in net flows. Given the lack of independent evidence on the magnitude of these costs, I choose the simpler specification in which the shock process is characterized by persistent innovations.

[13] In the post-WWII period the manufacturing industry has progressively moved away from the North-East and toward the South-West of the U.S., attracted by more favorable union and corporate legislation and fiscal incentives (see Peet (1983), Cobb (1993), and English (2006) for a historical perspective on this issue).

where c denotes consumption of goods and l^c represents consumption of services provided by land. The function ϕ is such that $\phi' > 0$, $\phi'' < 0$, and $\lim_{x \to 0} \phi'(x) \to \infty$. The multiplicative idiosyncratic shock υ summarizes a worker's match with the location in which she lives.[14] The shock takes one of two possible values, $\upsilon \in \left[\upsilon^b, \upsilon^g\right]$ where $0 \leq \upsilon^b < \upsilon^g$, with probabilities p and $1-p$ respectively. In subsequent periods the agent draws new match shocks. The latter evolve over time as a Markov process with transition probability $\pi^{\upsilon\upsilon'}$. Specifically, in the quantitative part of the paper I will assume that a bad match shock with a location remains bad forever ($\pi^{\upsilon^b\upsilon^b} = 1$), while a good shock remains good with probability δ ($\pi^{\upsilon^g\upsilon^g} = \delta$).[15] An agent's flow budget constraint in a period is:

$$w = c + rl^c,$$

where the the prices w and r refer to the location where the worker resides and there is no saving. Agents maximize the present value of (3), discounting future utility at the rate $\beta < 1$.

Timing. The sequence of events for an agent that has just arrived in a location at the end of period $t-1$ is as follows:

1. At the beginning of t, the agent draws the initial idiosyncratic match shock υ_t.

2. The agent optimally chooses c_t and l_t^c and receives a utility flow given by equation (3).

3. The agent observes the new value of υ in the location for period $t+1$, denoted by υ_{t+1}, and decides whether to stay in the same location or move to another location. If she decides to move she obtains expected utility e.

[14] In a previous version of the paper, I adopted Kennan and Walker (2006) specification and assumed that workers first moved into a location and then discovered their efficiency units of labor there. These two versions of the model produce nearly identical quantitative implications.

[15] Some exogenous depreciation of a good match is necessary to avoid that in the long-run all agents are characterized by good matches. While the model abstracts from explicit life-cycle considerations in accounting for migration decisions, the transition from a good to a bad idiosyncratic shock might be interpreted as the birth of a new generation of workers. This new generation will migrate in search of a good match until one is found. With this interpretation, migration rates would decline with age due to the selection effect associated with sampling from the density of the idiosyncratic match shock.

4. At the beginning of period $t+1$, if the agent remains in the same location in which she was living in t, she starts the period with the idiosyncratic shock υ_{t+1}. Otherwise, if the agent chooses to move to a new location, she draws a new initial idiosyncratic location-match υ as in step 1.

5. The same steps 2–4 are followed in period $t+1$. Thus, the decision to migrate or not is revisited by the worker in each period.

Search. Search is assumed to be directed toward the location that offers the highest expected utility, denoted by e. The timing of the model is such that an agent must decide whether to migrate or not from a certain location before the realization of next period's aggregate shock z' in that location. An agent that decides to migrate and has to determine where to direct herself is assumed to know only the expected realization of the shock z' in all potential locations of choice.

Value Function. At the beginning of a period the state of a location is fully characterized by the vector (y, n, ζ), where n denotes the measure of workers with match value υ^b. Denote the laws of motion for y and n by $Y(s,\zeta)$ and $N(s,\zeta)$ respectively, and summarize the two in the law of motion for $s=(y,n): s'=S(s,\zeta)$, where $S(s,\zeta)=[Y(s,\zeta), N(s,\zeta)]$. The dynamic programming problem of a worker characterized by idiosyncratic match υ with the location is given by:

$$\begin{aligned} V(s,\zeta,\upsilon) &= \max_{c,l^c} \left\{ \upsilon \times (c+\phi(l^c)) + \beta \sum_{\upsilon' \in \{\upsilon^b,\upsilon^g\}} \pi^{\upsilon\upsilon'} \max\left[\int V(s',\zeta',\upsilon')\, Q(\zeta, d\zeta'), e \right] \right\} \\ &\quad \text{s.t.} \qquad (4) \\ w(s,\zeta) &= c + r(s,\zeta)\, l^c, \\ s' &= S(s,\zeta). \end{aligned}$$

Inflows and Outflows. Denote gross inflows into a location characterized by state (s,ζ) by $x(s,\zeta)$ and gross outflows from that location by $o(s,\zeta)$. Let $q(s,\zeta,\upsilon')$ denote the

share of workers with idiosyncratic shock υ' who choose to leave a location characterized by state (s, ζ). Outflows from that location, denoted by $o(y, \zeta)$, are then equal to:

$$o(s, \zeta) = (n + (y - n)(1 - \delta))\, q(s, \zeta, \upsilon^b) + \delta\,(y - n)\, q(s, \zeta, \upsilon^g), \tag{5}$$

where the first and second terms on the right-hand side denote outflows by agents with bad and good shocks respectively.

Then, the laws of motion $Y(s, \zeta)$ and $N(s, \zeta)$ can be written as:

$$Y(s, \zeta) = y + x(s, \zeta) - o(s, \zeta), \tag{6}$$

$$N(s, \zeta) = (n + (y - n)(1 - \delta))\left(1 - q(s, \zeta, \upsilon^b)\right) + px(s, \zeta). \tag{7}$$

Equation (6) states that changes in the workforce equal the difference between inflows and outflows. Equation (7) gives the measure of workers characterized by idiosyncratic shock υ^b at the beginning of next period as the sum of two terms. The first term on the right-hand side of (7) represents the measure of workers with shock υ^b in the previous period who decided not to migrate. The second term represents the share of workers who migrates into the location and draws the shock υ^b.

Stationary Distribution. I consider a stationary environment with a location-invariant distribution of workers across locations $\mu(s, \zeta)$. The latter is such that:

$$\mu(S', \Xi') = \int_{\{s,\zeta : s' \in S'\}} Q(\zeta, \Xi')\mu(ds, d\zeta). \tag{8}$$

3.2 Equilibrium

A recursive stationary equilibrium for this economy is represented by a value function $V(s, \zeta, \upsilon)$, a migration probability $q(s, \zeta, \upsilon')$, laws of motion $S(s, \zeta) = [Y(s, \zeta), N(s, \zeta)]$, gross inflows $x(s, \zeta)$, gross outflows $o(s, \zeta)$, the expected utility of migration e, a station-

ary distribution $\mu(s,\zeta)$, demand for land by consumers $l^c(s,\zeta)$, demand for land by firms $l^f(s,\zeta)$, a wage function $w(s,\zeta)$, and a rent function $r(s,\zeta)$ such that:

- The value function $V(s,\zeta,v)$ satisfies the Bellman equation (4) given e and the law of motion $S(s,\zeta)$. In addition, the demand for land by consumers in location (s,ζ) satisfies the first order condition:

$$r(s,\zeta) = \phi'\left(l^c(s,\zeta)\right). \tag{9}$$

- The law of motion $S(s,\zeta)$ is related to $x(s,\zeta)$, $o(s,\zeta)$ and $q(s,\zeta,v')$ by equations (6)-(7).

- Inflows $x(s,\zeta)$ are consistent with directed search by migrating workers:

  - If $x(s,\zeta) = 0$ then:

$$e \geq p\int V\left(S(s,\zeta),\zeta',v^b\right)Q(\zeta,d\zeta') + (1-p)\int V\left(S(s,\zeta),\zeta',v^g\right)Q(\zeta,d\zeta').$$

  - If $x(s,\zeta) > 0$ then:

$$e = p\int V\left(S(s,\zeta),\zeta',v^b\right)Q(\zeta,d\zeta') + (1-p)\int V\left(S(s,\zeta),\zeta',v^g\right)Q(\zeta,d\zeta').$$

- Outflow probabilities $q(s,\zeta,v')$ are consistent with individual optimization:

  - If $q(s,\zeta,v') = 0$ then:

$$\int V\left(S(s,\zeta),\zeta',v'\right)Q(\zeta,d\zeta') \geq e,$$

  - If $q(s,\zeta,v') = 1$ then:

$$\int V\left(S(s,\zeta),\zeta',v'\right)Q(\zeta,d\zeta') \leq e,$$

– If $q(s, \zeta, \upsilon') \in (0, 1)$ then:

$$\int V\left(S\left(s, \zeta\right), \zeta', \upsilon'\right) Q(\zeta, d\zeta') = e.$$

- Aggregate population has measure one:

$$\int y\mu\left(ds, d\zeta\right) = 1.$$

- The invariant distribution $\mu\left(s, \zeta\right)$ is consistent with individual decisions, so equation (8) holds.

- The wage and rent functions in each location are such that:

$$\begin{aligned} r(s, \zeta) &= zF_l\left(y, l^f(s, \zeta)\right), \\ w(s, \zeta) &= zF_y\left(y, l^f(s, \zeta)\right). \end{aligned} \tag{10}$$

- The land market in each location clears:

$$yl^c(s, \zeta) + l^f(s, \zeta) = 1. \tag{11}$$

4 Empirical Implementation

This section describes the empirical implementation of the model. Most of the model's parameters are set a-priori while the ones governing the exogenous shock process are calibrated by targeting two moments of interest in the data.[16]

[16] When bringing the model to the data, one has to keep in mind that the model assumes the existence of a continuum of locations, and therefore a constant expected utility of migration e. The assumption of a continuum of location is mainly for feasibility: allowing for a finite number of locations in the theoretical model would make it virtually impossible to solve because a worker in a location would have to take into account the dynamics of the distribution of the state vector across all other locations when solving her dynamic programming problem.

A period in the model is taken to represent 1 year. The discount factor β is set equal to 0.96, implying a yearly discount rate of 4 percent. As anticipated in the previous section, I set $\pi^{\upsilon^g \upsilon^g} = \delta$ and $\pi^{\upsilon^b \upsilon^b} = 1$. Thus, an agent that draws a bad shock keeps that shock as long as she remains in the location, while an agent that draws a good shock might experience a transition to the bad state. An important implication of this formulation is that agents that draw a bad shock always choose to migrate because there is no cost of doing so and there is no possibility of a bad shock turning into a good one as long as the agents remains in the same location. The value of the bad shock υ^b is set to 0. The choice of υ^b implies that the value function of an agent with shock υ^b is simply $V\left(s, \zeta, \upsilon^b\right) = \beta e$. In turn, knowledge of this value function implies that the value function of an agent with shock υ^g is bounded in the interval $[e, e\,(1 - p\beta)\,/\,(1 - p)]$, where the upper bound represents the expected utility of an agent that draws a good shock υ^g in a location that experiences a positive inflow of workers ($x(s, \zeta) > 0$). As shown in Appendix C, under this specification of the model, the equilibrium can be computed by iterating on the recursive functional equation (22). The value of the good shock υ^g is set to one. This is just a normalization.

The parameters δ and p are set jointly to match the one and five-year migration rates in the economy's non-stochastic steady state:

$$\begin{aligned} \text{one-year migration rate} &= o = \frac{1-\delta}{1-p\delta}, \\ \text{five-year migration rate} &= o \times \left[1 + (1-p)\left(\delta + \delta^2 + \delta^3 + \delta^4\right)\right]. \end{aligned}$$

A lower value of δ and a higher value of p increase the migration rates. However, these two parameters play different roles. The parameter δ determines the migration propensity of workers who are incumbents in a location (i.e., had been living there for more than a period), while the parameter p determines the migration propensity of newly arrived workers. Thus, a higher value of p translates into a higher fraction of total migration accounted for by repeat-migrants, i.e., workers who move for at least two periods in a row. In turn, given o,

the higher the incidence of repeat-migration (the higher p), the lower the five-year migration rate is. The 1-year migration rate from the American Community Survey of 2000 is about 2.2 percent, while the average five-year migration rate across Census years is about 8.1 percent. In order to target these two migration rates the parameters δ and p are set at 0.98 and 0.29 respectively.

The production technology and the utility function are assumed to take the following forms:

$$F(y,l) = y^{\tau} l^{1-\tau}, \ \tau \in (0,1), \tag{12}$$

$$\phi(l) = \frac{A}{\alpha} l^{\alpha}, \ \alpha \in (-\infty, 1), \ A > 0. \tag{13}$$

The parameter τ represents the share of labor income in total output. To calibrate this parameter, it is necessary to take into account the fact that physical capital has already been solved out of the profit optimization problem of the firm. Ciccone (2002) reports an income share of land in manufacturing of 1.5 percent. Based on this information, the parameter τ is set equal to 0.985. The elasticity parameter α is set equal to $1-\tau$. This value greatly facilitates the numerical solution of the model because it implies that land can be solved out of the model analytically (see Appendix C). The parameter A is set to match the share of land relative to wage income in the steady state. Roback (1982) finds that the budget share of land in the data is about 6 percent. Taking into account that labor income is about 2/3 of total income, this yields a value of $A = 0.086$.

In order to obtain a stationary process for $\{z\}$ it is necessary that the parameter ψ be strictly smaller than one. The value of the parameter ψ is set exogenously equal to 0.999. This value is such that the growth rate of the shock process z is approximately an AR(1), while at the same time preserving the stationarity of z. The parameters ρ and σ_u are set by matching some of the cross-sectional and time-series moments of states' net inflow rates

reported in Section 2.2. Specifically, the parameter σ_u is identified by the standard deviation of net inflow rates constructed by pooling all the adjusted state-year data together. The value of this standard deviation is 3.2 percent (Table 2). The parameter ρ is identified by the first-order autocorrelation coefficient of net inflow rates across Census years, which is 0.67 (Table 5). Notice that if the level of the productivity shock z, rather than its growth rate ε, had followed an AR(1) process, the first order autocorrelation coefficient of net flow rates would have been close to zero, which is strongly at odds with the data. Table 7 summarizes the calibrated values of the model's parameters.[17]

5 Model's Fit

In this section I present the main quantitative predictions of the model. In what follows, I emphasize in bold the moments that are targeted in the calibration of the model. The first dimension along which I evaluate the model's fit is the relationship between net flow rates, labor earnings, and land rents (Table 8).

Both in the model and in the data, the cross-sectional correlation between net flow rates and average earnings and rents is quite small. The intuition for this result is that net flow rates depend on the *growth rate* of total factor productivity in a location, ε, while wages and rents depend on the *level* of productivity, z. In locations with higher levels of productivity land rents and wages must be higher than in locations where productivity is low to make both firms and workers indifferent with respect to their choice of location (Roback (1982)). Thus, wages and rents are positively correlated in the cross-section. Moreover, the rank correlation coefficients between state-level earnings and rents in a Census year and the equivalent measures in the previous Census year are positive and large both in the data and in the model, suggesting considerable persistence in states' rankings.

Table 9 reports the cross-sectional dispersion of worker flows across locations, as measured by their standard deviation. The model accounts for the ranking of dispersions observed in

[17] Appendix D describes in detail the numerical algorithm used to solve and calibrate the model.

the data, with gross inflow rates being more dispersed than net flow rates and the latter more dispersed than gross outflow rates. This ranking implies that gross inflow and outflow rates are positively correlated and so are gross and net inflow rates, as shown in Table 10.

Figure 4(a) plots gross outflow against gross inflow rates using data generated by the benchmark version of the model. As Table 10 and Figure 4(a) show, the model can account quite well for the positive correlation between gross inflow and outflow rates. The positive correlation between inflow and outflow rates is due to a selection effect, as incoming workers have a higher propensity to migrate from the location to which they have arrived than workers already living there. Upon arrival incoming workers draw their idiosyncratic match with the location from the unconditional distribution of v. A fraction $p \approx 0.3$ of these will draw a bad idiosyncratic match and choose to migrate again at the end of the period. Instead, the idiosyncratic probability of migration for pre-existing residents of the location is equal to $1 - \delta \approx 0.015$, the rate at which good matches depreciate.

Note however that the model generates too little dispersion in outflow rates (Table 9). This is due to two reasons, further explored in Section 6. First, the parameter p is relatively small, implying that only one in three incoming workers chooses to migrate again. A higher value of p generates more outmigration in response to a positive gross inflow. Second, in the benchmark calibration of the model, well-matched workers tend *not* to leave locations hit by negative productivity shocks. As a consequence, the correlation between net flows and gross outflows, which is close to zero in the data, is positive in the model (see Figure 5(a) with a quadratic fit to the data). More specifically, in the data the relationship between net flows and gross outflows is U-shaped. The calibrated model accounts for the increasing portion of the U. In order to explain the decreasing portion of the U as well, it would have to be the case that as net flow rates become negative a location adjusts to further declines in productivity by means of higher gross outflows. In this region of the state space, some of the workers who have a good idiosyncratic match with the location would choose to out-migrate. It turns out that in the benchmark calibration workers with good idiosyncratic matches rarely choose

to migrate, for two reasons. First, a location tends to exogenously lose population due to the depreciation of good matches. This outflow is large enough that no further adjustments are needed in the local labor market. Second, upon moving well-matched workers face a relatively large probability of drawing a bad match in the new location. This acts as an implicit cost of migration. Section 6 shows how the model can be made consistent with the outflow-net flow correlation by increasing the variance of the productivity shock u'.

Finally, Table 11 represents the first-order autocorrelation coefficients of net and gross worker flows. The model generates persistent worker flows, even if the autocorrelation of gross inflow rates is smaller than in the data.

6 Sensitivity Analysis and Extensions

In this section I perform some sensitivity analysis in order to provide intuition for the features of the model that are essential to generate the main results. For sake of simplicity I focus on a subset of the moments of interest, summarized in Table 12.

Sensitivity Analysis. I start by modifying some of the model's parameters, one at a time, while keeping the others fixed at their benchmark values.

- *Higher probability of bad idiosyncratic match.* The correlation between inflow and outflow rates and the cross-sectional dispersion of outflow rates both increase with p, the probability of drawing a bad shock upon migration. As an example, I consider the case $p = 0.5$. Figure 4(b) plots inflow and outflow rates for this case.

- *Higher standard deviation of location-specific shocks.* A feature of the benchmark calibration of the model is that net flow rates and gross outflow rates are positively correlated, while in the data this correlation is close to zero (Tables 3 and 4). As mentioned in the previous section this is due to the fact that well-matched workers rarely choose to leave a location, even when its productivity is declining. One way to make this channel quantitatively more important is to increase the variance of the

productivity shock u'. Figure 5(b) represents the scatterplot between net flow rates and gross outflow rates corresponding to the case when σ_u is four times larger than in the benchmark calibration.[18] The figure also shows the U-shaped quadratic fit of this relationship. The correlation between these two variables is smaller than in the benchmark and close to zero.

- *No persistence in the growth rate of productivity shocks.* In this case ($\rho = 0$), net flow rates are basically uncorrelated over time.

- *No ex-post heterogeneity among agents.* In this version of the model ($\upsilon^g = \upsilon^b$), gross and net inflows coincide when the latter are positive while gross outflows and absolute net flows coincide when net flows are negative. Hence the negative correlation between gross inflows and outflow rates.

Extensions. I consider two extensions of the benchmark model.

- *Mobility Costs.* The basic model can readily accommodate mobility costs. Some authors - such as Kennan and Walker (2006) - have estimated large moving costs across U.S. states. Letting k denote such costs, the qualitative properties of the benchmark model are the same as long as in equilibrium $e > k$, i.e., the expected utility of migration is larger than the cost of migration. Table 12 presents results for the case $k = 4.8$, which corresponds to five times the wage income in the model's steady state. This figure is equivalent to about \$220,000 in 2005 and is comparable with the estimates of mobility costs by Kennan and Walker (2006). As in the benchmark also in this case badly matched workers will always choose to move independently of the value of k. However, the presence of a migration cost reduces the migration incentives of workers with good matches when their location is hit by negative productivity shocks. Since in the benchmark version of the model, these workers rarely choose to move, the presence of a mobility cost does not significantly affect the key moments of interest.

[18] Note that the parameter σ_u in this exercise is four times larger than in Table 7, even if due to approximation when comparing Tables 7 and 12 it appears to be five times larger.

- *Non-Directed Search.* The benchmark model assumes directed search. It is instructive to consider the opposite extreme according to which migrating workers are assigned to locations with equal probability. Thus, moving workers cannot arbitrage away differences in expected utility across locations and each location receives a constant measure of workers x. In the benchmark calibration of the model, v^g workers never choose to migrate out of a location. Thus, the constant inflow x must be balanced in every period by a corresponding constant outflow of badly matched workers. In equilibrium all locations have constant population of one. Local prices fluctuate in response to productivity shocks, but gross worker flows are the same across locations and net flows are zero.

7 Additional Evidence on the Mechanism

A key prediction of the model is that states with large inflow rates in one period should experience relatively large outflow rates in *subsequent* periods, rather than contemporaneously. In other words, the observed positive contemporaneous correlation between gross inflow and outflow rates documented in Tables 3 and 4 should really be interpreted as the by-product of the positive correlation between current outflows and lagged inflows combined with the high persistence displayed by gross inflow rates (Table 5). I can evaluate the extent to which this argument is correct by using the time-series dimension of worker flow data across several Census decades. Ideally, one would want to regress the outflow rates between years $T-5$ and T on both the contemporaneous inflow rates between $T-5$ and T and the lagged inflow rates between $T-10$ and $T-5$. In order for the data to be consistent with the model, the current outflow rate should display a higher correlation with the *lagged* inflow rate than with the current inflow rate. Unfortunately, given that Census data are only available once every decade, the second-best option is to use, as proxy for lagged inflows, the inflows that occurred between year $T-15$ and year $T-10$, instead of the ones between $T-10$ and $T-5$.

The results of this exercise are reported in Table 13.

Within a given state, a higher-than-average *lagged* inflow rate of workers is positively associated with the current outflow rate, even controlling for the *current* inflow rate. The latter displays a negative partial correlation with the contemporaneous outflow rate. This evidence is consistent with the mechanisms highlighted in this paper.

8 Summary and Conclusions

This paper considers a general equilibrium model of gross and net migration across locations. The calibrated model is consistent with many of the cross-sectional and time-series facts about worker flows and state-level labor earnings and rents. In particular, the model can account for the fact that in the data U.S. states with relative high inflow rates of workers are also characterized by relatively high outflow rates. Standard models of employment reallocation across geographic areas cannot account for this pattern.

The main mechanism emphasized in the theoretical model is unobserved heterogeneity among otherwise identical workers. The model generates ex-post differences among workers due to matching effects between a worker and a location. An implication of the paper is that cross-sectional heterogeneity in out-migration rates across U.S. states cannot be explained exclusively by the observable sources of heterogeneity among workers, such as age and education, traditionally emphasized in the migration literature. It also cannot be accounted for exclusively by aggregate shocks to locations, as emphasized in the macro-labor literature. Instead, the model generates dispersion in outflow rates across states through the interaction of location-specific shocks and ex-post heterogeneity among workers. From this perspective, this paper has made progress toward bridging the gap between the mostly partial-equilibrium literature on migration (e.g., Kennan and Walker (2006)) and the general equilibrium approach to employment reallocation which typically does not allow for workers' idiosyncratic shocks (e.g. Lucas and Prescott (1974)).

Decomposing net flows into their gross components sheds some light on how some states adjust to economic shocks. In response to a positive productivity shock a state absorbs internal migrants while losing a substantial portion of these newcomers in subsequent periods. States with high employment growth are characterized by relatively large gross outflow rates. Moreover, the growth rate of local labor demand shocks appears to be very persistent over time, as certain locations experience positive employment growth for several decades before the effect of a shock dies out. While the model presented here is purely descriptive, one can view the exogenous driving shock z as being affected by a state's local development policies. An implication of this broader interpretation of z is that, due to diminishing returns to capital and labor in the production function, one-time permanent changes in local policies do not generate sustained employment growth at the state level. After an initial growth of employment, a location simply ends up with a permanently higher level of employment, wages and rents. An interesting but challenging task would be to determine empirically whether increasing returns at the local level, due for example to agglomeration effects, might generate persistent employment growth in a location in response to one-time shocks to local labor demand.

A few other extensions of the basic setting presented here are worth mentioning. First, in the model there are only two possible levels for the idiosyncratic shock, good and bad. To be able to keep the model computationally tractable I calibrate it in such a way that badly-matched agents always choose to migrate, and only well-matched ones face the trade-off of whether to stay and keep the good idiosyncratic match or leave and attain a better wage-rent combination. A richer but computationally intractable version would allow badly matched agents to migrate in certain states of the world and stay in others. In this case, I would expect locations hit by declining productivity shocks to shed labor by means of higher outflows, a channel that is not active in the benchmark calibration of the model.

Second, it would be useful to dig deeper into the nature of idiosyncratic migration decisions. In this paper the latter stem from a match shock between a worker and a location.

This is a convenient reduced-form way of capturing a variety of factors from true individual taste shocks about a location that are uncovered only after a move has been made to information about the nature of one's job. Of course, in real life not all information is revealed after a move and individuals sample job offers from firms in other locations without having to move. This alternative setting might account for the positive inflow-outflow correlation if job offers across state borders are concentrated on a subset of workers who are ex-ante, as opposed to ex-post, different from the rest of the population and therefore behave as repeat-movers. Given the cross-sectional nature of the Census micro data employed in this paper, it is very difficult to distinguish between the model with ex-post heterogeneity among workers and a version with ex-ante differences.

Another interesting, although challenging, extension would be to allow for return migration, which is the focus of Kennan and Walker (2006), and explore its impact on the aggregate dynamics of gross worker flows.

I leave these extensions for future research.

References

Alvarez, Fernando and Marcelo Veracierto (2000): "Labor-Market Policies in an Equilibrium Search Model," in NBER Macroeconomics Annual, Volume 14. Cambridge and London: MIT Press, 265-304.

Alvarez, Fernando and Marcelo Veracierto (2006): "Fixed-Term Employment Contracts in an Equilibrium Search Model," NBER working paper series no. 12791.

Armenter, Roc and Francesc Ortega (2007): "Credible Redistributive Policies and Migration Across U.S. States," Working paper Federal Reserve Bank of New York.

Bartik Timothy (1991): *Who Benefits from State and Local Development Policies?*, W.E. Upjohn Institute for Employment Research, Kalamazoo Michigan.

Blanchard, Olivier and Lawrence Katz (1992): "Regional Evolutions," *Brooking Papers on Economic Activity* 1992, 1-75.

Card, David (2001): "Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration," *Journal of Labor Economics* 19, 22-64.

Caselli, Francesco and Wilbur Coleman (2001): "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy* 109, 584-616.

Clark, Todd (1998): "Employment Fluctuations in U.S. Regions and Industries: The Roles of National, Region-Specific, and Industry-Specific Shocks," *Journal of Labor Economics*, 202-229.

Cobb, James (1993): The Selling of the South: the Southern Crusade for Industrial Development 1936-1990, University of Illinois Press, 2nd edition.

Davis Steven, Prakash Loungani, and Ramamohan Mahidhara (1991): "Regional Unemployment Cycles," Working Paper, University of Chicago.

English, Beth (2006): A Common Thread: Labor, Politics, and Capital Mobility in the Textile Industry, Georgia University Press.

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

Greenwood, Michael (1975): "Research on Internal Migration in the United States: A Survey," *Journal of Economic Literature* 13, 397-433.

Greenwood, Michael and Gary Hunt (1984): "Migration and Interregional Employment Redistribution in the United States," *American Economic Review* 74, 957-969.

Kennan, John and James Walker (2006): "The Effect of Expected Income on Individual Migration Decisions, " Working Paper, University of Wisconsin-Madison.

Kambourov, Gueorgui and Iourii Manovskii (2009): "Occupational Mobility and Wage Inequality," *Review of Economic Studies* 76, 731-759.

Jovanovic, Boyan (1979): "Job Matching and the Theory of Turnover," *Journal of Political Economy* 87, 972-990.

Jovanovic, Boyan and Robert Moffitt (1990): "An Estimate of a Sectoral Model of Labor Mobility," *Journal of Political Economy* 98, 827-852.

Lkhagvasuren, Damba (2005): "Big Locational Differences in Unemployment Despite High Labor Mobility," Working Paper, University of Rochester.

Lucas, Robert and Edward Prescott (1974): "Equilibrium Search and Unemployment," *Journal of Economic Theory* 7, 188-209.

Miller, Ann (1967): "The Migration of Employed Persons to and from Metropolitan Areas of the United States," *Journal of the American Statistical Association* 62, 1418-1432.

Miller, Robert (1984): "Job Matching and Occupational Choice," *Journal of Political Economy* 92, 1086-1120.

Peet, Richard (1983): "Relations of Production and Relocation of United States Manufacturing Industry Since 1960," *Economic Geography* 59, 112–143.

Press W., Teukolsky, S., Vetterling, W., Flannery, B., 1996. Numerical Recipes in Fortran 77 (Cambridge University Press, Cambridge).

Rappaport J. (2004): "Why are Population Flows so Persistent?," *Journal of Urban Economics* 56, 554–580.

Roback, Jennifer (1982): "Wages, Rents and the Quality of Life," *Journal of Political Econ-*

omy 90, 1257-1278.

Steven Ruggles, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander (2004). Integrated Public Use Microdata Series: Version 3.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor].

Sjaastad, Larry (1961): "Migration in the Upper MidWest," in *Four Papers on Methodology*, Upper Midwest Economic Study, Technical Paper no.1, 41-69.

Sjaastad, Larry (1962): "The Costs and Returns of Human Migration," *Journal of Political Economy* 70, 80-93.

Topel, Robert (1986): "Local Labor Markets," *Journal of Political Economy* 94 (Supplement), S111-43.

Treyz, George, Rickman, Dan, Hunt Gary and Michael Greenwood (1993): "The Dynamics of U.S. Internal Migration," *Review of Economics and Statistics* 75, 209-214.

Van Nieuwerburgh, Stjin and Pierre-Olivier Weill (2006): "Why Has House Price Dispersion Gone Up?" NBER working paper series no. 12538.

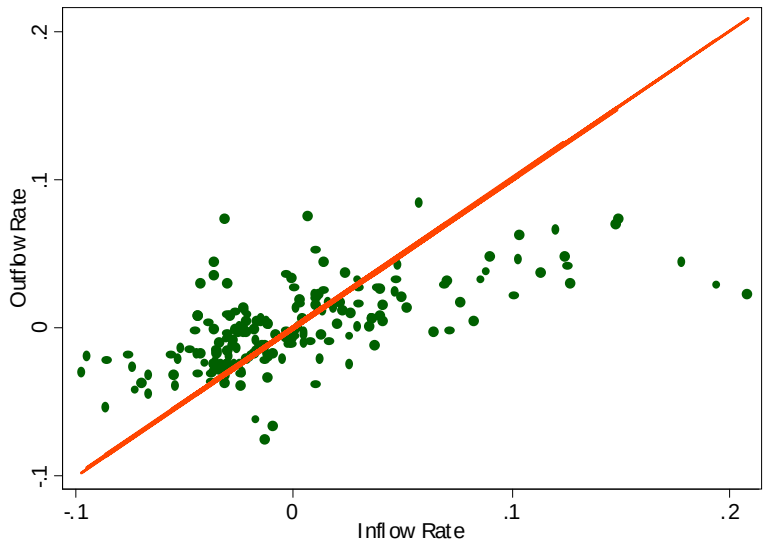

Figure 1: Scatter plot of inflow and outflow rates of workers across U.S. states, 1970-2000. The figure also reports the 45 degree line. Data source: my computations based on data from the U.S. Census. See Appendix A.2 for details on the construction of inflow and outflow rates.

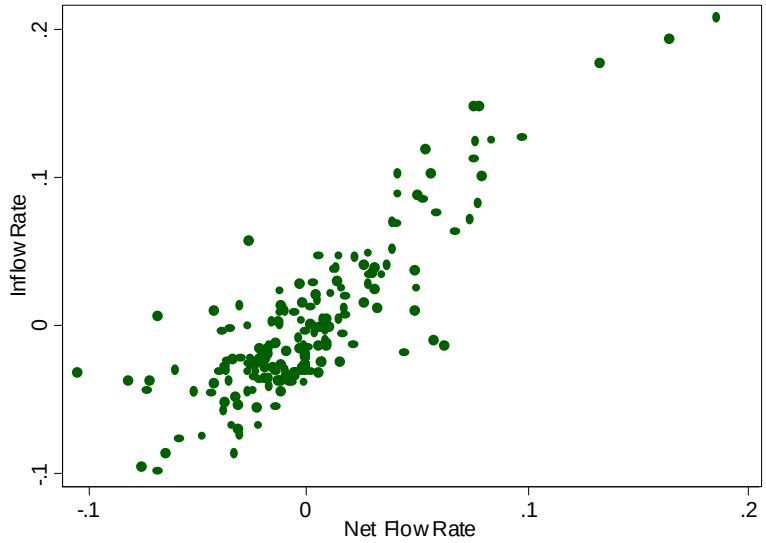

Figure 2: Scatter plot of net and gross inflow rates of workers across U.S. states, 1970-2000. Data source: my computations based on data from the U.S. Census.

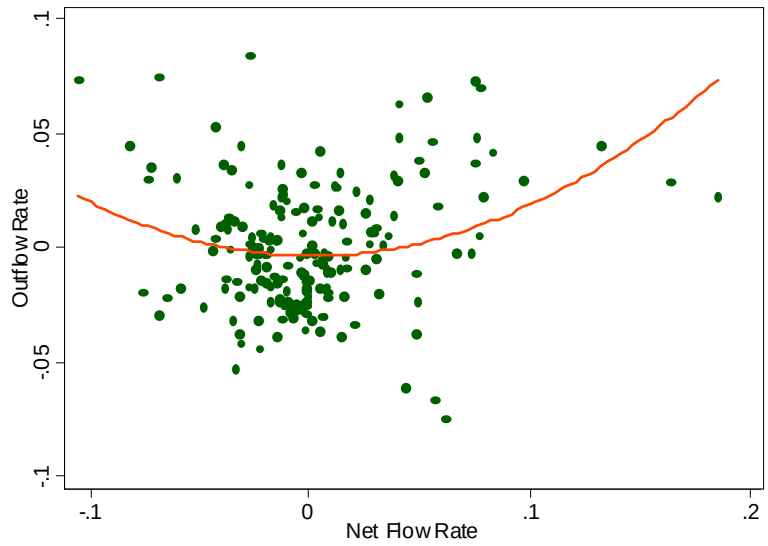

Figure 3: Scatter plot of net flow and gross outflow rates of workers across U.S. states, 1970-2000. Data source: my computations based on data from the U.S. Census.

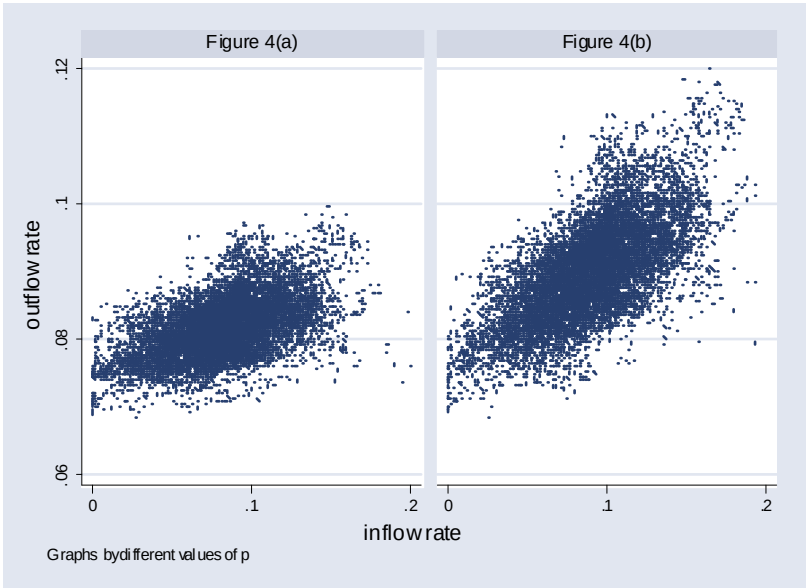

Figure 4: Scatter plots of gross outflow rates against gross inflow rates using data generated by the model (10,000 observations). Figure 4(a) corresponds to the benchmark calibration of the model. Figure 4(b) corresponds to the case in which $p = 0.5$ and all remaining parameters remain set at their benchmark values.

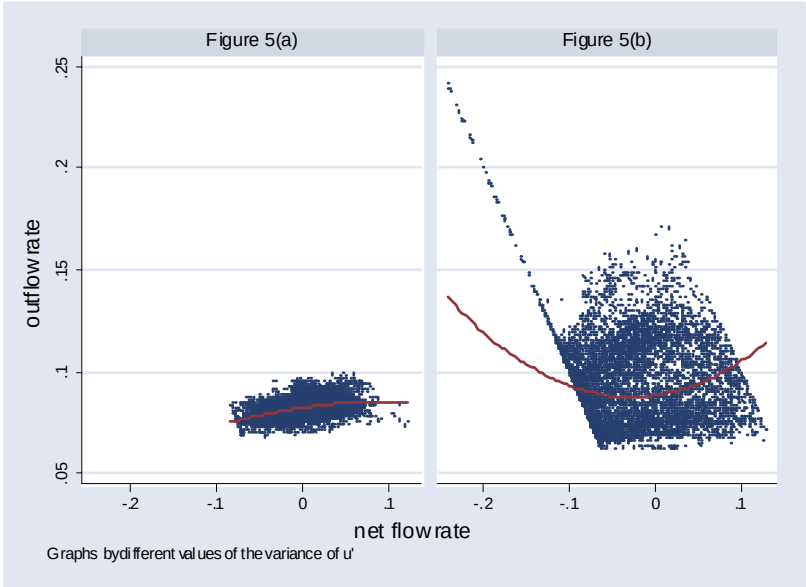

Figure 5: Scatter plots of gross outflow rates against net flow rates using data generated by the model (10,000 observations). Figure 5(a) corresponds to the benchmark calibration of the model. Figure 5(b) corresponds to the case in which σ_u is four times larger than in the benchmark and all remaining parameters remain set at their benchmark values. A quadratic fit to the data is shown in both figures.

Table 1

Magnitude of Worker Flows Across U.S. States

Census Year	Gross Worker Reallocation	Average Absolute Net Flow Rates	Excess Worker Reallocation
1970	13.40	2.35	11.05
1980	17.26	3.34	13.92
1990	17.62	3.44	14.18
2000	16.80	2.13	14.67
Average	16.27	2.81	13.46

Table 2

Standard Deviations of Worker Flows

Outflow Rate	Inflow Rate	Net Flow Rate
2.25	4.44	3.20

Table 3

Correlation Coefficients Among Worker Flows

	Outflow Rate	Inflow Rate
Net Flow Rate	0.31**	0.88**
Inflow Rate	0.73**	

Note: ** significant at 1% level.

Table 4

Within-Group Correlations Among Worker Flows

	Outflow Rate			Inflow Rate		
Percentiles	*25th*	*50th*	*75th*	*25th*	*50th*	*75th*
Net Flow Rate	-0.39^{**}	-0.17	-0.04	0.76^{**}	0.80^{**}	0.85^{**}
Inflow Rate	0.28^{**}	0.45^{**}	0.55^{**}			

Note: ** significant at 1% level. * significant at 5% level.

Table 5

First-Order Autocorrelation of Worker Flows

Net Flow Rate	Outflow Rate	Inflow Rate
0.67^{**}	0.81^{**}	0.93^{**}

Note: ** significant at 1% level.

Table 6

State-Level Earnings, Rents, and Net Flow Rates

	Cross-Sectional Correlations		Spearman Rank
	Rents	Net Flow Rates	Correlation with Previous Census
Earnings	0.80^{**}	0.02	0.79^{**}
Rents		0.04	0.77^{**}

Note: ** significant at 1% level. * significant at 5% level.

Table 7

The Model's Parameters

	Parameter	Value
β	discount factor	0.9600
δ	transition probability of idiosyncratic shock	0.9845
τ	production function parameter	0.9850
α	utility function parameter	0.0150
A	utility function parameter	0.0863
υ^b	value of bad idiosyncratic shock	0.0000
υ^g	value of good idiosyncratic shock	1.0000
p	initial probability of drawing bad shock υ^b	0.2979
ψ	mean-reversion parameter	0.9990
ρ	persistence of labor demand shock	0.9772
σ_u	volatility of labor demand shock	0.0001

Table 8

State-Level Earnings, Rents, and Net Flow Rates

	Cross-Sectional Correlations				Spearman	
	Rents		Net Flow Rates		Rank Correlation	
	Data	Model	Data	Model	Data	Model
Earnings	0.80	0.99	0.02	0.18	0.79	0.99
Rents			0.04	0.20	0.77	0.99

Table 9

Standard Deviations of Worker Flows

Outflow Rate		Inflow Rate		Net Flow Rate	
Data	Model	Data	Model	Data	Model
2.25	0.005	4.44	3.41	**3.20**	**3.20**

Table 10

Correlation Coefficients Among Worker Flows

	Outflow Rate		Inflow Rate	
	Data	Model	Data	Model
Net Flow Rate	0.31	0.41	0.88	0.99
Inflow Rate	0.73	0.52		

Table 11

First-Order Autocorrelations of Worker Flows

Net Flow Rate		Outflow Rate		Inflow Rate	
Data	Model	Data	Model	Data	Model
0.67	**0.67**	0.81	0.86	0.93	0.68

Table 12

Comparisons Across Model Economies

	Average (%)	Std. Dev. (%)	Cross-Sectional Correlations			Persistence
	O	N	I-O	N-O	N-I	N
Data	8.13	3.20	0.73	0.31	0.88	0.67
Benchmark	8.13	3.20	0.52	0.41	0.99	0.67
Sensitivity						
$p = 0.5$	8.94	3.22	0.67	0.53	0.98	0.68
$\sigma_u = 0.0005$	8.23	10.68	0.34	0.18	0.98	0.66
$\rho = 0$	8.13	0.003	0.16	-0.07	0.97	0.04
$\upsilon^g = \upsilon^b = 1$	1.25	1.75	-0.52	-0.87	0.88	0.77
Extensions						
Mobility Costs	8.13	3.20	0.52	0.41	0.99	0.67
Random Search	8.13	n.a.	n.a.	n.a.	n.a.	n.a.

Legend: N=net flow rate, O=outflow rate, I=inflow rate, n.a. due to constant worker flows

Table 13

Current Outflows and Lagged Inflows

Dependent variable: Gross outflow rate$_{T-5,T}$	
Gross inflow rate$_{T-5,T}$	-0.19^{**}
Gross inflow rate$_{T-15,T-10}$	0.45^{**}
Year dummies	yes
State dummies	yes
Number of observations	144
Adjusted R^2	0.64

Note: ** significant at 1% level.

NOTE: APPENDICES B, C, AND D ARE NOT MEANT FOR PUBLICATION

A Data Appendix

A.1 Sample Selection

Data are from the five percent samples of the 1980-2000 Censuses and from the one percent sample of the 1970 Census (Form 1 State Sample). All the measures of gross and net flows and the stock of population that are reported in the paper are computed using a sample of individuals that, at the time of the relevant Census, satisfy the following restrictions:

- are between 28 and 60 years of age;
- are not living in group quarters;
- are in the labor force but not in the armed forces;
- if foreign-born, have immigrated to the U.S. at least five years before the Census year;
- were not living abroad five years before the Census year;
- are not living in either Alaska, Hawaii, or the District of Columbia in the Census year or five years before the Census year.

A.2 Construction of Flow Rates

For each Census year the raw data are adjusted in a series of steps to control for composition effects and gravity forces.

Composition effects. I regress an individual i's stay-move choice on a set of dummies for education, age, sector, race, occupation, number of children, marital status, sex, and immigration status. Denote by m_i individual i's residual from this linear probability model.

Each individual is observed living in a state k at the time of the Census and in a state j five years before the Census year. The residuals are then aggregated up by state of origin-state of destination pairs to create a matrix of bilateral gross flows among the 48 states:

$$F_{jk} = \sum_{i \in (j,k)} m_i,$$

where the notation $i \in (j,k)$ means "individual i living in state k in the Census year and in state j five years earlier." The flows are then converted into bilateral outflow and inflow *rates* by dividing by a state's population five years before the Census year (P):

$$\text{outflow}_{jk} = \frac{F_{jk}}{P_j}, \ \text{inflow}_{jk} = \frac{F_{jk}}{P_k}.$$

Gravity. States might display certain bilateral patterns of inflows and outflows due to their relative size and their distance from one another (the "gravity forces"). In order to control for gravity, I separately regress the bilateral inflow and outflow rates defined above on log distance, log distance squared, log population of origin and destination states, and the squares of these two variables, a border dummy, and the interactions between the border dummy and log distance, log population of the origin state, log population of the destination state. The residuals of these two equations for bilateral inflow and outflow rates are then aggregated into an outflow and inflow rate for each of the 48 states. The net flow rate for a state is defined as the difference between the gross inflow and outflow rate.

A.3 Demographic Groups

In order to compute the within-group correlations in Table 4, I construct 385 demographic groups based on the following variables and data from the 2000 Census:

- age (`age`); 7 age groups: 28-31, 32-36, 37-41, 42-46, 47-51, 52-56, 57-60;
- education (`educ99`); 5 education groups: high-school dropout, high-school diploma,

some college, college degree, above college;

- industry of employment (`ind1990`); 11 industries: (1) agriculture, fishing, forestry, hunting and mining, (2) construction, (3) manufacturing non-durables, (4) manufacturing durables, (5) transportation, communication and other public utilities, (6) wholesale and retail trade, (7) finance, insurance, and real estate, (8) business and repair services, (9) personal services, entertainment and recreation services, (10) professional and related services, (11) public administration.

For each of the 385 demographic groups I then construct outflow, inflow and net flow rates by dividing the worker flows by the group's population living in a given state five years prior to the Census year. These group-specific rates are used to compute the statistics in Table 4.

A.4 Weekly Earnings

Workers' weekly earnings are computed using data from the Census of Population and Housing. The same sample selection criteria listed above in Section A.1 are also applied in this case. Weekly earnings are obtained by summing, for each worker, annual wage income (`incwage`) and business and farm income (`incbus+incfarm`), and dividing the sum by the number of weeks worked (`wkswork1`). Each source of income refers to the year preceding the Census year. I have dropped from the sample a very small number of observations for which an individual reported zero annual earnings but a positive number of weeks worked. In a few instances, reported earnings by self-employed individuals are negative, and these observations have been dropped as well. Given that earnings refer to the year prior to the Census and the worker's labor force participation status refers to the time of the survey, a small fraction of individuals (about 2.5 percent of the sample) reported zero annual earnings and zero weeks worked in the year prior to the Census. I have also dropped these individuals from the sample.

For each Census year, the logarithm of weekly earnings was regressed on the following variables: 48 dummies for workers' state of residence in the Census year (`statefip`), a measure of workers' experience (computed subtracting years of education from the workers' age) and experience squared, 17 education dummies (`educ99`), a workers' sex (`sex`), 3 race dummies ("white", "black" and "others", constructed from `raced`), 11 sectoral dummies (constructed from `ind1990`), and 6 occupational dummies (constructed from `occ1990`). The R^2 of these regressions was typically 30 percent.

The measure of average weekly earnings for each state is represented by the estimates of state fixed-effects in this regression.

A.5 Rents

The Census of Population and Housing provides data on the gross monthly rent (`rentgrs`) paid by a renter. This variable reports the gross monthly rental cost of the housing unit, including contract rent plus additional costs for utilities (water, electricity, gas) and fuels (oil, coal, kerosene, wood, etc.). This information is used to derive a measure of land rents in each U.S. state. Observations on rents are obtained for those workers renting a housing units who satisfy the sample selection criteria listed above in Section A.1. In each Census year, about 30 percent of the sample obtained from applying those selection criteria rents (as opposed to owns) its housing unit. For example, in the 2000 Census more than one million observations are available for renters.

In order to remove the influence of observable characteristics of a housing unit from the monthly rent, I run a hedonic regression of the logarithm of the rent on state fixed effects and a list of observable characteristics of the housing units. These include dummies for: whether the housing unit is located in a metropolitan area (`metro`), whether the unit is used commercially (`commuse`), the acreage of the property (`acreprop`), the acreage of the house (`acrehous`), whether meals are included in the rent (`rentmeal`), whether the housing unit is in a condominium (`condo`), whether the housing unit contains a kitchen (`kitchen`), the

number of rooms (`rooms`), the availability of plumbing facilities (`plumbing`), the age of the unit (`builtyr`), the number of bedrooms (`bedrooms`), the number of units in the structure (`unitsstr`).

The measure of average rents at the state level is represented by the estimated state fixed-effects from this regression.

A.6 Immigration

The sample of workers selected according to the criteria spelled out in Section A.1 includes foreign-born workers provided they migrated to the U.S. at least five years before the Census year. In what follows I refer to these workers as "recent" immigrants. While this selection guarantees that aggregate net flows of workers across U.S. states are zero in each Census year, it also assumes away recent immigration flows. The latter might potentially play an important role in affecting internal migration flows from and into certain states. In order to quantify these effects and to place the magnitude of internal migration flows in perspective, relative to their international counterpart, I have used the 2000 Census data to compute for each U.S. state the ratio between the number of recent immigrants that are located in that state in the year 2000 and the state's 1995 population. The data indicate that the average rate of recent immigration across U.S. states is about 1.5 percent, while Table 1 indicates that the average inflow rate of workers from the rest of the U.S. in the 2000 Census is about 8.4 percent. Thus, for the average state internal migration is much larger than recent immigration. Of course, it is well known that recent immigrants tend to cluster in a few locations. There are three states for which the recent immigration rate is at least 30 percent as large as the gross inflow rate due to internal migration: California (47 percent), New York (56 percent), and New Jersey (34 percent). It turns out, however, that for the cross-section of U.S. states there is no significant association between rates of recent immigration and both gross and net flow rates due to internal migration. So, there is no evidence supporting the view that states with low gross inflow rates experienced larger than average immigration

from abroad. This result is consistent with the findings of Card (2001) for U.S. metropolitan areas.

B Extension: Amenities in the Utility Function (NOT FOR PUBLICATION)

It is easy to modify the model to include location-specific amenities and to show that an appropriate choice of functional forms yields exactly the same implications for net and gross flows of labor as the one derived in the previous section. Suppose that there are no location-specific productivity shocks z, but rather that each location is characterized by a time-varying amenity a. The utility function takes the form:

$$u = \upsilon \times h(a) \times (c + \phi(l^c)),$$

where $h(.)$ is an increasing function of a.

Solving for the optimal choice of land and consumption yields the following indirect utility function:

$$u\left(s, \widehat{\zeta}, \upsilon\right) = \upsilon h(a)\left(w\left(s, \widehat{\zeta}\right) - r\left(s, \widehat{\zeta}\right) l^c\left(s, \widehat{\zeta}\right) + \phi\left(l^c\left(s, \widehat{\zeta}\right)\right)\right),$$

where $\widehat{\zeta}$ denotes the vector (a, ε).

Replacing the expressions for the wage and rent yields:

$$u\left(s, \widehat{\zeta}, \upsilon\right) = \upsilon h(a)\, w\left(s, \widehat{\zeta}\right)\left(1 + \left(\frac{A}{1-\tau}\right)^{\frac{1}{\tau}}\right).$$

Since wages are given by:

$$w\left(s,\widehat{\zeta}\right)=\frac{\tau y^{\tau-1}}{\left[1+\left(\frac{A}{1-\tau}\right)^{\frac{1}{\tau}}\right]^{1-\tau}},$$

I obtain that

$$u\left(s,\widehat{\zeta},\upsilon\right)=\upsilon h\left(a\right)\tau y^{\tau-1}\left(1+\left(\frac{A}{1-\tau}\right)^{\frac{1}{\tau}}\right)^{\tau}.$$

This model has the same reduced-form utility function as the benchmark model (equation 16) provided that $h(a)$ takes the functional form:

$$h(a)=\frac{\left(a^{\frac{1}{\tau}}+\left(\frac{A}{1-\tau}\right)^{\frac{1}{\tau}}\right)^{\tau}}{\left(1+\left(\frac{A}{1-\tau}\right)^{\frac{1}{\tau}}\right)^{\tau}},$$

and a follows the same stochastic process as z:

$$\begin{aligned} a' &= a\varepsilon', \\ \varepsilon' &= a^{\psi-1}\varepsilon^{\rho}u'. \end{aligned}$$

Notice that the model with amenities has the same implications for flows of workers across states, but has different implications for prices of labor and land. Specifically, rents would tend to be higher and wages lower in locations with better amenities (as in Roback (1982)).

C The Benchmark Model (NOT FOR PUBLICATION)

First of all, notice that workers whose match is bad ($\upsilon = \upsilon^b$) will always choose to migrate because migration is costless and search is directed. Thus, $q(s,\zeta,\upsilon^b)=1$. Given the

parameterization $\upsilon^b = 0$, their value function is then given by:

$$V\left(s, \zeta, \upsilon^b\right) = \beta e. \tag{14}$$

Consider now the agents with a good idiosyncratic shock. Instantaneous utility is defined as:

$$u\left(s, \zeta, \upsilon\right) = \upsilon \times \left(w\left(s, \zeta\right) - r\left(s, \zeta\right) l^c\left(s, \zeta\right) + \phi\left(l^c\left(s, \zeta\right)\right)\right).$$

It is possible to simplify the expression for $u\left(s, \zeta, \upsilon\right)$ using the functional form for ϕ in equation (13), the condition $\alpha = 1 - \tau$, and the fact that due to the constant returns to scale assumption unit cost of production must equal one (the price of output). The latter equation, which can be easily obtained by manipulating the firms' first order condition, can then be used to solve for $r\left(s, \zeta\right)$ as a function of $w\left(s, \zeta\right)$ and the shock z:

$$r\left(s, \zeta\right) = \left[\frac{z\left(1-\tau\right)^{1-\tau}\left(\tau\right)^{\tau}}{w\left(s, \zeta\right)^{\tau}}\right]^{\frac{1}{1-\tau}}. \tag{15}$$

The indirect utility function $u\left(s, \zeta, \upsilon\right)$ becomes:

$$u\left(s, \zeta, \upsilon\right) = \upsilon w\left(s, \zeta\right)\left(1 + z^{-\frac{1}{\tau}}\left(\frac{A}{1-\tau}\right)^{\frac{1}{\tau}}\right). \tag{16}$$

Land can be completely solved for by using the equilibrium conditions (10) and (11) together with the worker's first order condition for land, equation (9). Replace the latter into equation (11) to get:

$$y\left(\frac{A}{r\left(s, \zeta\right)}\right)^{\frac{1}{1-\alpha}} + l^c\left(s, \zeta\right) = 1.$$

Now, solve for $l^c\left(s, \zeta\right)$ and replace it in the inverse demand function for labor:

$$w\left(s, \zeta\right) = \tau z y^{\tau-1}\left(1 - y\left(\frac{A}{r\left(s, \zeta\right)}\right)^{\frac{1}{1-\alpha}}\right)^{1-\tau}.$$

Use (15) to replace $r(s,\zeta)$ on the right-hand side of this equation. This yields an equation with $w(s,\zeta)$ on both the right and left-hand sides. The advantage of the assumption $\alpha = 1-\tau$ is that this equation can be solved in closed-form to yield the wage as a function of the state variables of the model:

$$w(s,\zeta) = \frac{\tau z y^{\tau-1}}{\left[1 + z^{-\frac{1}{\tau}} \left(\frac{A}{1-\tau}\right)^{\frac{1}{\tau}}\right]^{1-\tau}}. \tag{17}$$

I have then shown that it is possible to solve out for $r(s,\zeta)$ and land from the original specification and rewrite the model as one in which the utility functions take the form (16) and the unit wage is given by (17).

With these preliminary steps, the model's equilibrium can then be characterized in relation to the behavior of workers for whom $\upsilon = \upsilon^g$. We need to distinguish among three different regions of the state space:

- **Region 1**. Some (or all) of the υ^g workers choose to leave and some (or none) choose to remain. In this case, the value of staying must be less or equal to the value of leaving, with strict inequality prevailing if everybody chooses to leave:

$$\int V(s',\zeta',\upsilon^g)\, Q(\zeta, d\zeta') \leq e. \tag{18}$$

  Notice that in this case, endogenous inflows are zero, $x(s,\zeta) = 0$, because workers with a good idiosyncratic shock choose to leave and the expected match shock of incoming workers would be strictly less than υ^g. Thus, the components of s' are:

$$\begin{aligned} Y(s,\zeta) &= \delta(y-n)\left(1 - q(s,\zeta,\upsilon^g)\right), \\ N(s,\zeta) &= 0 \end{aligned}$$

Notice that $q(s,\zeta,\upsilon^g) \leq 1$ is implicitly defined by (18) in case of equality.

- **Region 2**. None of the υ^g workers chooses to leave ($q(s,\zeta,\upsilon^g) = 0$) and no new worker

chooses to locate there. In this case:

$$\int V\left(s',\zeta',\upsilon^g\right)Q(\zeta,d\zeta') > e, \tag{19}$$

$$p\int V\left(s',\zeta',\upsilon^b\right)Q(\zeta,d\zeta') + (1-p)\int V\left(s',\zeta',\upsilon^g\right)Q(\zeta,d\zeta') < e. \tag{20}$$

The first inequality expresses the fact that it is better for a υ^g type of worker to remain in the location, while the second inequality expresses the fact that no migrating worker will choose to migrate to this location. The endogenous components of s' are:

$$\begin{aligned} Y(s,\zeta) &= \delta(y-n), \\ N(s,\zeta) &= 0. \end{aligned}$$

Notice that, using equation (14), the two inequalities (19) and (20) can be rewritten as:

$$e < \int V\left(s',\zeta',\upsilon^g\right)Q(\zeta,d\zeta') < e\left(\frac{1-p\beta}{1-p}\right).$$

- **Region 3**. None of the υ^g workers chooses to leave and some new workers choose to locate there. In this case, equation (19) still holds, while equation (20) holds as an equality:

$$e < \int V\left(s',\zeta',\upsilon^g\right)Q(\zeta,d\zeta') = e\left(\frac{1-p\beta}{1-p}\right). \tag{21}$$

The endogenous components of s' are:

$$\begin{aligned} Y(s,\zeta) &= x(\zeta,s) + \delta(y-n), \\ N(s,\zeta) &= px(\zeta,s). \end{aligned}$$

where $x(s,\zeta)$ is implicitly determined by the equality condition in (21).

It follows from the discussion above that the value function for an agent with a good

match υ^g satisfies the following functional equation:

$$\begin{aligned} V(s,\zeta,\upsilon^g) &= \upsilon^g(c+\phi(l^c)) + \\ &\quad \beta\delta \min\left\{\max\left[\int V(S(s,\zeta),\zeta',\upsilon^g)\,Q(\zeta,d\zeta'),e\right], e\left(\frac{1-p\beta}{1-p}\right)\right\} + \beta(1-\delta)e, \end{aligned} \tag{22}$$

where the endogenous components of s' are:

$$\begin{aligned} Y(s,\zeta) &= \delta(y-n), \\ N(s,\zeta) &= 0. \end{aligned}$$

D Details On Numerical Implementation (NOT FOR PUBLICATION)

This section describes the steps that I follow in solving and calibrating the model. The algorithm is comprised of three loops: one for finding the value function conditional on e and the parameter vector (ρ,σ_u), one for finding the equilibrium value of e for given (ρ,σ_u), and one for finding (ρ,σ_u) in order to match the empirical moments of interest. Every change in (ρ,σ_u) entails new equilibrium values for e, while a new guess for e requires the computation of the associated value function.

Step 1 (Guess). Start from an initial guess for the parameter vector (ρ,σ_u) and for the value of search e. The guess for e is updated in Step 3 below, while the guess for (ρ,σ_u) is updated in Step 4.

Step 2 (Dynamic Programming). Solve the dynamic programming problem for υ^g workers by iterating on the functional equation (22). This is the most time-consuming step of the procedure because there are four continuous state variables in the problem (recall that s includes y and n while z includes ζ and ε) and because the procedure involves numerical integration of the value function with respect to the density of the innovation u'. The

solution of the dynamic programming problem yields gross inflows $x(s, \zeta)$ and the probability of outflow $q(s, \zeta, v^g)$ as functions of the state vector (s, ζ). These two functions allow one to recover all the other variables of interest conditional on (s, ζ).

Step 3 (Equilibrium). Solve for the equilibrium value of e by defining the function $f(e)$:

$$f(e) = \int y\mu(ds, d\zeta) - 1, \tag{23}$$

The integral in equation (23) is computed by simulating the economy for a very large number of periods (6 million) and computing the average population:

$$\frac{1}{T}\sum_{t=1}^{T} y_t.$$

The real number e^* such that $f(e^*) = 0$ is computed using Brent's method for finding the zero of a function.

Step 4 (Calibration). Given e^*, it is possible to compute the equilibrium value of all the variables of interest. The vector (ρ, σ_u) is estimated by constructing the model counterpart of the two moments listed in the text (Section 4) and choosing (ρ, σ_u) so that the model-generated moments are exactly equal to their empirical counterparts. Since there are two parameters and two moments, this is an exactly identified model. The problem then becomes one of solving two non-linear equations in two unknowns. The model-generated moments are constructed by using 5 million simulated data drawn from the model. Each simulated moment is then compared with its empirical counterpart. In order to find a solution for this non-linear system of two equations in two unknowns I have used Broyden's algorithm. The latter operates in the following way (for a more detailed description, see Press et al. (1996, chapter 9)). First, it numerically approximates the Jacobian matrix associated with the non-linear system. It then uses this approximate Jacobian to find an updated vector (ρ, σ_u) by implementing the Newton step, which guarantees quadratic convergence if the initial guess is close to the solution. If the Newton step is not "successful", the algorithm

tries a smaller step by backtracking along the Newton dimension. When an acceptable step is determined, (ρ, σ_u) is updated and the algorithm can proceed in the way described above, once an updated Jacobian has been obtained. Since the numerical computation of the Jacobian can be costly (and in this model it is), the Jacobian at the new vector (ρ, σ_u) is iteratively approximated using Broyden's formula. The non-linear solver stops when the maximum percentage difference between the simulated moments and the empirical moments is, in absolute value, smaller than 10^{-4}.