# Understanding Gross Workers Flows Across U.S. States

Daniele Coen-Pirani*
Tepper School of Business
Carnegie Mellon University

This Version: May 2006. First Version: February 2006.
PRELIMINARY AND INCOMPLETE - COMMENTS WELCOME

**Abstract**

This paper documents and provides an explanation for the main stylized facts about net and gross workers flows across states in the U.S. While it is generally known that gross flows of population across locations are significantly larger in the U.S. than within most European countries, there is considerable heterogeneity in gross and net flows across states within the U.S. itself. The main purpose of the paper is to test whether a simple general equilibrium search model based on Lucas and Prescott (1974)'s island economy can account for the main stylized facts. The key stylized facts are as follows. In the cross-sectional dimension: (1) Gross inflow rates are more dispersed than net inflow rates, which are more dispersed than gross outflow rates. (2) Gross inflow and outflow rates are positively correlated. (3) Gross and net inflow rates are highly positively correlated, while net flow rates and gross outflow rates are uncorrelated. In the time-series dimension, there is a large degree of persistence in both gross and net flow rates across Census years for a given state. To address these facts, I develop a general equilibrium model of net and gross workers' flows across locations. Net flows are driven by shocks to local labor demand, while gross flows are driven by idiosyncratic location-specific shocks to workers' productivity. In response to shocks to the growth rate of labor productivity in a location, the model generates artificial data that is generally consistent with the stylized facts listed above.

*E-mail: coenp@andrew.cmu.edu. Thanks, without implication, to seminar participants at UT Austin, SUNY Albany, Cleveland Fed, and the 2006 Midwest Economics Association Meetings, and to Yoonsoo Lee, Holger Sieg, Fallaw Sowell, Marla Ripoll, Maria Ferreyra, Dennis Epple, Alexis Leon, Iourii Manovski, and especially Rui Castro, for useful discussions and suggestions. Richard Lowery provided able research assistance.

# 1 Introduction

This paper documents and provides an explanation for the main stylized facts about net and gross workers flows across states in the U.S. While it is generally known that gross flows of population across locations are significantly larger in the U.S. than within most European countries (see Hassler et al., 2005), there is considerable heterogeneity in gross and net flows across locations within the U.S. itself.[1] The main purpose of the paper is to test whether a simple general equilibrium search model based on Lucas and Prescott (1974)'s island economy can account for the main stylized facts.

I start by documenting these facts using the decennial Census of the U.S. for the post-WWII period. The latter allows one to determine a respondent's state of residence in the Census year as well as five years before the Census year. This information is used to construct state-level aggregate gross and net flow rates of workers. These flows are adjusted to take into account the different demographic and industrial composition of the workforce across states and differences in other state characteristics, such as size.

The key stylized facts are as follows. In the cross-sectional dimension: (1) Gross inflow rates are more dispersed than net inflow rates, which are more dispersed than gross outflow rates. (2) Gross inflow and outflow rates are positively correlated. (3) Gross and net inflow rates are highly positively correlated, while net flow rates and gross outflow rates tend to be uncorrelated. These facts seem to suggest that reallocation of population within the U.S. occurs mainly through variations in gross inflows (large in fast-growing states and small in slow-growing states), rather than in gross outflows. In other words, states that tend to lose population to other states do so by attracting fewer new workers as opposed to losing more local ones. In the time-series dimension, there is a large degree of persistence in both gross and net flow rates across Census years for a given state. Last, there is a significantly positive correlation between average state wages, adjusted by differences in living costs, and net flow rates.

In order to account for these facts, I consider a simple search model. The model economy is composed by a set of local labor markets ("islands"), that are hit by idiosyncratic labor demand shocks. Local wages would tend to rise in response to these shocks, but workers' mobility across islands equalizes the price of an efficiency unit of labor across islands. At a point in time, a location experiences both gross inflows and gross outflows. This is because a worker's idiosyncratic productivity differs across islands, feeding workers' gross flows. In general equilibrium, the value of migration is pinned down by a zero excess demand condition for aggregate net flows.

Preliminary results show that the model is consistent with the main stylized facts mentioned above. The mechanics of the model can be better understood by considering an unanticipated shock to local labor demand. On impact, the workers' net flow rate rises while the outflow rate remains constant. Afterwards, outflows rise above their steady state value, as some of the location's newly arrived workers are ex-post unlucky and decide to move again. The persistent nature of local labor demand shocks implies that net flows remain above steady state for several periods. Due to the response of gross outflows, gross inflows exceed net inflows. Thus, gross inflows are more volatile than net inflows, which, in turn, are more volatile than gross outflows. Gross inflow and outflow rates are positively correlated as larger gross inflows of workers lead to larger gross outflows.

This paper is related to several literatures. The closest literature is the one initiated by Lucas and Prescott (1974) in their "island" model of the labor market. Lucas and Prescott develop a model of workers' net flows across locations driven by shocks to local labor demand. In a sense the

[1]Internal migration rates for Italy and Spain in 1995 were about 1/4 of those for the U.S. In France and Germany they were respectively 67% and 55% of the U.S. values. See the OECD (2000) for details.

present paper can be thought of as a version of Lucas and Prescott (1974) in which also workers are hit by idiosyncratic location-specific productivity shocks, giving rise to gross flows of workers across locations. The introduction of these idiosyncratic shocks has implications for net flows as well. Under some conditions, the latter are such that unit wages are always equalized across locations despite the existence of mobility costs. In the original contribution by Lucas and Prescott, instead, the fact that workers are homogeneous within a location implies that wage differentials across locations are necessary to give rise to net flows.[2] The paper also builds on the contribution by Blanchard and Katz (1992), who developed a reduced form model of workers' net flows across U.S. states, and provide some interesting VAR evidence on the nature of states' adjustment process to local labor demand shocks. Relative to Blanchard and Katz, this paper also focuses on gross flows of workers.[3] Finally, this paper is related to the partial equilibrium literature on the determinants of workers' migration decisions. Kennan and Walker (2005) carefully estimate such model using NLSY data and use the panel structure of the data to identify wage differences due to location effects.[4]

The rest of the paper is organized as follows. Section 2 describes the data and the stylized facts. Section 3 describes the model. Section 4 offers a discussion of some modelling issues. Section 5 evaluates the model in light of the stylized facts of Section 2. Section 6 discusses possible extensions of the paper. Section 7 concludes. The data appendix offers a more detailed description of the data and the construction of the flow variables.

## 2 Data and Stylized Facts

**Data on Workers' Flows.** The main data set I use is the U.S. Census of Population for several decades.[5] The Census data have the clear advantage of being a large and comprehensive dataset. Information on geographic mobility of individuals is available from other sources. For example, the March Current Population Survey (March CPS) contains such information, but only includes approximately 60,000 households. Given that, on average only 3 percent of the population leaves its state of residence in a given year, this amounts to observing less than 2,000 households migrating across state lines, or, on average, 40 households per state. In contrast, the decennial Census typically contains information on million of households.

Since 1940, the Census questionnaire has included a question regarding the location (state and metropolitan area) where an individual was living 5 years before the Census interview. Using this information, I construct rates of gross and net flows of population across the 48 contiguous United States.[6] The population flows always refer to the 5 year period preceding the Census year, and represent a lower bound on the actual flows, as some individuals moved more than once during

[2] Topel (1986) considers a setting similar to Lucas and Prescott (1974).

[3] More recent contributions by macroeconomists to the literature on internal migration of workers include Hassler et al (2005) and Lkhagvasuren (2005). The former argue that differences in the generosity of unemployment insurance between the U.S. and Europe can explain higher internal mobility rates in the U.S. The latter paper tries to explain the existence of persistent differentials in unemployment rates across U.S. states by means of a general equilibrium matching model with location-specific idiosyncratic productivity shocks. The paper is also related to the traditional research on the determinants of population flows within the U.S., surveyed by Greenwood (1975) and more recently developed by Greenwood and Hunt (1984) and Treyz et al. (1993). The contribution of this paper relatively to this mostly empirical literature is to develop a tractable structural model of gross workers flows.

[4] See also Dahl (2002) and Borjas, Bronars and Trejo (1992).

[5] This is available online at `www.ipums.org/usa`.

[6] The levels of inflow and outflow of population for a given state were standardized by the number of individuals who were surveyed in the Census year and reported living in that state 5 years before. Net flow rates were defined as the difference between gross inflow and outflow rates.

these 5 years. In order to focus on geographic mobility that is not motivated by college attendance or retirement, I restrict attention to individuals who were between 27 and 60 years of age and in the labor force at the time of the Census. The appendix contains more detailed information on issues of sample selection as well as on the construction of the variables described below. From now on, for simplicity, I will refer to a state's "population" as the collection of individuals satisfying the sample selection criteria described in the appendix.

Before proceeding it is necessary to briefly comment on the choice of U.S. states as primary units of analysis. Since the focus of the paper is the *geographical* mobility of workers, the ideal unit of analysis should be a local labor market. The latter concept is intuitive but not simple to define unambiguously. In practice, a local labor market is often associated with a metropolitan area. In this paper I have chosen not to take a metropolitan area as the basic unit of analysis for several reasons. First, the 1970 Census does not report information on an individual's metropolitan area of residence in 1965. This information is instead available at the state level.[7] This is important because the information contained in the 1970-2000 Censuses is used below to estimate the stochastic process for local labor demand shocks. The lack of the 1970 data would further reduce the already short time-series dimension of the data. Second, about 20 percent of the U.S. population does not currently live in a metropolitan area. This figure has increased by about 10 percentage points since 1970, and it displays a non-trivial geographic variation. Therefore, also in this case there would be some ambiguity associated with the definition of a local labor market. Third, according to the Census there are more than 200 metropolitan areas in the U.S. This figure makes the estimation of the model extremely lengthy, while it is feasible, yet long, to work with 48 locations. Last, for the purpose of policy analysis, labor market policies (e.g. unemployment insurance) are set at the state level.

**Composition Effects and Heterogeneity Across States.** Figures 1-3 report scatter plots of outflow, inflow and net flow rates computed using the raw data from the 2000 Census. There is, of course, considerable heterogeneity among states in at least two dimensions. First, at the micro level, different states have a different composition of population, in terms of age, education, industry of employment, etc. When comparing measures of population flows across states, one has to make sure to control for possible composition effects. It can, in fact, be that certain states exhibit higher gross flows because of the sectoral or demographic composition of their employment structure. For example, if the gambling industry has a particularly high turnover of workers', then we might expect the state of Nevada, in which this industry is particularly large, to feature large inflows and outflows of workers. To address the micro heterogeneity, I divide the population into 490 demographic groups defined by age, education, and industry. Then, I compute gross outflow and inflow rates for each state and for each demographic group. Last, I compute the state-wide rates as a weighted average of the groups' rates, using as a weight for each group its relative size in the U.S. population. It turns out that the gross flows obtained using this procedure are very close to the unadjusted ones.[8] Thus, composition effects due to cross-state heterogeneity in the age, education and industry affiliation of the states' population do not seem to play an important role in explaining differential gross population flows across states.

At the macro level, states have different sizes, different numbers of large metropolitan areas, etc. The concern here is that differences in gross flows might be driven by some of these factors, as opposed to the economic forces I would like to emphasize. To address this macro heterogeneity, I have run a cross-sectional regression of inflow and outflow rates (adjusted using the above procedure)

[7]Specifically, the Census variable `migmet5` (metropolitan area of residence 5 years before) is not available in 1970, while the variable `migplac5` (state of residence 5 years before) is.

[8]The cross-sectional correlation between adjusted and unadjusted rates in the 2000 Census is always above 0.97.

on states' land area, number of metropolitan areas with population above 0.5 million, and year when the state formally joined the U.S. I have then defined the outflow and inflow rates to be the residuals of this regression, and the net flow rates as the difference between the two. This second adjustment has a more sizeable effect on the statistics of interest, but does not affect the basic properties of the data, either.

**Inflows and Outflows in the Cross-Section.** Tables 1 and 2 below provides descriptive statistics regarding inflow, outflow and net flow rates across U.S. states using data from the Census 2000, adjusted as described above.[9]

Table 1
Basic Statistics on Workers Flows (Census 2000)

| | Mean | Median | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|---|
| Outflow Rate | 8.86 | 8.67 | 1.54 | 5.46 (Wisconsin) | 17.15 (Wyoming) |
| Inflow Rate | 8.86 | 8.27 | 3.33 | 4.03 (North Dakota) | 28.26 (Nevada) |
| Net Flow Rate | 0.00 | −0.10 | 2.59 | −7.99 (North Dakota) | 12.51 (Nevada) |

Table 2
Cross-Sectional Correlations (Census 2000)

| | Outflow Rate | Inflow Rate | Net Flow Rate |
|---|---|---|---|
| Outflow Rate | 1 | 0.66*** | 0.25* |
| Inflow Rate | | 1 | 0.89*** |
| Net Flow Rate | | | 1 |

*** denotes significant at 1% level, * at 10% level

From these two Tables, some interesting facts emerge:

- There is a relatively large dispersion across states in outflow, inflow and net flow rates, with outflow rates being relatively less dispersed than gross and net inflow rates (Table 1).
- States that experience a relative large gross inflow of population also tend to experience a relatively large gross outflow of population (Table 2 and Figure 1). For example, the state of Nevada ranked first in terms of gross inflows (about 28.26 percent) and second in terms of gross outflows (15.74 percent). Interestingly, the positive correlation between gross inflows and outflows is apparently a well-known, though not extensively documented, stylized fact in the literature on internal migration of population (see Greenwood, 1975).[10] The correlations in Table 2, particularly the one between gross inflows and gross outflows, are consistent with two polar situations. One in which the large flows are symptoms of a changing industry/demographic mix of the state's workforce, so that the outgoing workers are different from the incoming ones. The other in which the "same" type of worker moves in and out of the state. In order to distinguish between these two possibilities, I have used the 490 demographic groups (indexed by $g$ and described above) and computed, for each state $j$ and for

[9] The statistics in this and the following tables are computed weighting each state by its relative population.

[10] See, for example, Miller (1967, page 1426, Table 3). She defines locations in terms of metropolitan areas, instead of states and shows, using 1960 Census data, that this correlation is robustly positive both in the aggregate and within demographic groups defined by sex, race and occupational category.

the 2000 Census, the following measure of within demographic groups workers' geographical reallocation:[11]

$$\frac{\sum_g (in_{jg} + out_{jg}) - \sum_g |in_{jg} - out_{jg}|}{\sum_g (in_{jg} + out_{jg}) - \left|\sum_g (in_{jg} - out_{jg})\right|}. \quad (1)$$

The denominator of this expression gives the difference between the sum of gross inflow and outflow from location $j$ and the absolute net flow. Thus, it represents the excess of workers' mobility over and above what is needed to accommodate net workers' flows. To understand the numerator of equation (1), suppose that inflows and outflows of workers always occurred between demographic groups. This means that for each group $g$, we would either have $in_{jg} > 0$ and $out_{jg} = 0$ or $in_{jg} = 0$ and $out_{jg} > 0$. In this case the numerator of (1) would be zero, and so would the measure of within-group reallocation. At the other extreme, if inflows and outflows were always balanced within groups ($in_{jg} = out_{jg}$), then the index would be equal to one. The outflows-weighted average of this measure across states for the 2000 Census was 0.91, suggesting that most flows occur within the demographic/industry groups described above. A way to consider exclusively within-group flows when computing the correlations of Table 2, is to compute the cross-sectional correlation between gross inflows and outflows for each group separately. Then, the 490 correlation coefficients can be averaged using as weights the groups' population shares in the U.S. The following Table reports these adjusted correlation:

Table 3
Adjusted Cross-Sectional Correlations (Census 2000)

| | Outflow Rate | Inflow Rate | Net Flow Rate |
|---|---|---|---|
| Outflow Rate | 1 | 0.39 | −0.23 |
| Inflow Rate | | 1 | 0.79 |
| Net Flow Rate | | | 1 |

Comparing the correlations in Tables 2 and 3, one notices that the latter are smaller. The signs of the correlations between gross inflows and outflows are the same in both Tables, while the correlation between outflows and net flows turns negative in Table 3. These results are consistent with the view that some of the gross flows we observe have to do with changes in the composition of states' workforce. From our perspective, however, it is important that, even within narrowly defined demographic/industry groups, there is a positive correlation between gross inflows and outflows.[12]

- Gross and net inflows are highly and positively correlated in the cross-section, while the correlation between net flows and gross outflows is in absolute value smaller (Tables 2 and 3 and Figures 2 and 3). This observation, together with the previous two, seems to suggest that reallocation of population within the U.S. occurs mainly through variations in gross inflows (large in fast-growing states and small in slow-growing states) rather than in gross outflows across states. In other words, states that tend to lose population to other states seem to do so by attracting fewer new workers as opposed to losing more local ones.

[11] This measure has been used, for example, by Davis and Haltiwanger (1992) to decompose aggregate excess reallocation of jobs into a between-sector and a within-sector component.

[12] For consistency, one could also adjust the cross-sectional standard deviation of flow rates in Table 1 in order to capture exclusively the dispersion of flows within demographic groups, and not the cross-groups covariance terms. By doing so, since the latter covariances tend to be negative, one would obtain higher standard deviations for inflow, outflow, and net flow rate. Their ranking does not change, though. For simplicity, I do not carry out this further adjustment.

**Cross-Sections Over Time and the Time-Series Dimension of Workers' Flows.** It is natural to ask whether the statistics presented in the previous Tables are peculiar to the 2000 Census or not. It is also important to determine how much persistence there is in gross population flows for a given state. Both questions can be answered by considering other Census years.

Tables 4 and 5 confirm that the salient features of gross and net flows pointed out above in relation to the 2000 Census are also present in the 1970-1990 Censuses.[13],[14]

Table 4
Basic Statistics on Population Flows (Censuses 1970-2000)

| | 2000 | | 1990 | | 1980 | | 1970 | |
|---|---|---|---|---|---|---|---|---|
| | Mean | St. Dev. | Mean | St. Dev. | Mean | St. Dev. | Mean | St. Dev. |
| Outflow Rate | 8.59 | 1.50 | 9.04 | 2.01 | 8.97 | 1.77 | 6.89 | 1.72 |
| Inflow Rate | 8.59 | 3.22 | 9.04 | 3.90 | 8.97 | 3.86 | 6.89 | 2.73 |
| Net Flow Rate | 0.00 | 2.60 | 0.00 | 3.84 | 0.00 | 3.06 | 0.00 | 2.39 |

Table 5
Cross-Sectional Correlations (Censuses 1970-1990)

| | Outflow Rate | Inflow Rate | Net Flow Rate |
|---|---|---|---|
| Outflow Rate | | | |
| 1970 | 1 | 0.50*** | -0.15 |
| 1980 | 1 | 0.63*** | 0.22 |
| 1990 | 1 | 0.29** | -0.23 |
| 2000 | 1 | 0.60*** | 0.17 |
| Inflow Rate | | | |
| 1970 | | 1 | 0.78*** |
| 1980 | | 1 | 0.89*** |
| 1990 | | 1 | 0.86*** |
| 2000 | | 1 | 0.89*** |

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

Do states with relatively high gross and net flows between 1995 and 2000, also tend to display relatively high flows between 1985 and 1990, and before? The answer to this question is affirmative for both gross and net flows. The following Table reports, for each type of flow, its autocorrelation coefficient across Census years, computed by pooling all state-year data points together.[15]

[13]Extending the analysis before 1970 presents some difficulty. The 1960 Census does not report a person's state of residence in 1955, but only if the person migrated across states or not. Thus, in 1960 it is only possible to compute gross inflows, but not gross outflows or net flows. In the 1950 Census, the migration question pertains to one year before, rather than 5 years before. I exploit the 1950 Census year in Table 6 below. The 1940 Census does not present particular problems.

[14]For simplicity, given that the first type of adjustment mentioned above (for composition effects related to age, education and industry of employment) did not produce any sizeable effect on the statistics of Tables 1 and 2, in this Table the data for 1970-2000 are only subject to the second type of adjustment mentioned above. This explains the difference between the results in Table 1 and the one in this Table for the 2000 Census.

[15]In order to increase the sample size in the time-series dimension, I have included the 1950 Census in these computations. The 1950 Census asked respondents to report their state of residence in 1949, as opposed to 1945. The 1 year migration data were converted into 5 year migration data by multiplying the 1 year flows by 4.538. This number guarantees that the cross-sectional standard deviation of net flow rates in 1950 is the same as the average standard deviation in all the previous Census years. Notice that in Table 6 the difference between $t$ and $t-1$ is equivalent to 10 years.

Table 6
Autocorrelations of Population Flows (Censuses 1950, 1970-2000)

| | Autocorrelation Coefficient | | | | |
|---|---|---|---|---|---|
| | $t, t-1$ | $t, t-2$ | $t, t-3$ | $t, t-4$ | $t, t-5$ |
| Outflow Rate | 0.66*** | 0.68*** | 0.29*** | 0.49*** | 0.59*** |
| Inflow Rate | 0.86*** | 0.71*** | 0.56*** | 0.73*** | 0.73*** |
| Net Flow Rate | 0.73*** | 0.61*** | 0.55*** | 0.41*** | 0.40*** |

*** significant at 1% level.

Notice that all these flow rates are very persistent over time.

**Migration Motives.** Not all moves of workers across states are motivated by economic reasons. The Census survey does not contain any question regarding a worker's reason for the change of residence. However, since 1999, the March CPS has asked this question. I have aggregated the different answers to this question into two categories, according to whether an interstate move is attributable or not to economic factors. The appendix gives more information about the criteria used for this assignment. About two thirds of all interstate moves that occurred between 1999 and 2003, according to the March CPS, were of the first kind.

**Earnings.** Using the 2000 Census I have constructed a measure of a worker's weekly earnings in 1999. These nominal figures were deflated using the ACCRA cost of living index, which measures the relative price of a given basket of goods and services in a number of U.S. cities.[16] Denote the real weekly earnings of worker $i$ in 1999 by $w_i$. The Census data provides detailed information regarding a worker's demographic characteristics, occupation and industry. Let these observables be summarized by the vector $x_i$. In order to construct measures of average wages within states and of residual wage inequality, I have run the following regression for log earnings:

$$\ln w_i = \mu \times STATE_i + \omega \times MOVE_i + x_i\beta + \varepsilon_i, \tag{2}$$

where $STATE_i$ is a vector of dummy variables recording individual's $i$ state of residence in 2000, while $MOVE_i$ is a dummy variable that equals one if worker $i$ moved across state lines sometimes between 1995 and 2000 and zero otherwise.[17] The estimates of this regression reveal two interesting facts. First, the estimates of the dummy coefficients $\mu$ reveal that, controlling for observables, the standard deviation of weekly earnings across states is about 11 percent of average weekly earnings in the U.S.[18] Second, the estimate of $\omega$ in equation (2) suggests that the weekly earnings of workers

[16] See the appendix for more detail on this index. Due to limitations in the availability of this index, the size of the sample used to carry out the computations of this section is about 50 percent of the original sample size used in previous sections. However, it still includes about 2.5 million observations.

[17] As Kennan and Walker (2005) point out, the residual $\varepsilon_i$ can be interpreted as reflecting the influence of (at least) three orthogonal factors on a worker's earnings: a location match component (which varies across locations for the same worker, but is constant within locations), a worker's fixed effect (which is constant within and across locations for a given worker), and a transitory effect (which varies both within and across locations for a given worker). Location-specific effects on earnings are a key determinant of migration decisions, in addition to differences in average earnings across states. Of course, the cross-sectional nature of the Census data does not allow one to separately identify the three components of residual wages. Fortunately, Kennan and Walker (2005, page 11) provide such decomposition using NLSY data. According to their estimates, location effects explain about 16 percent of the overall inequality in residual earnings among young high school educated workers.

[18] This figure is calculated as $std\left\{\exp\left(\widehat{\mu}_j\right)\right\}/E\left\{\exp\left(\widehat{\mu}_j\right)\right\}$. Notice that in computing this number I am imposing the same vector of observables in all states.

that have moved across state lines in the 5 years preceding the 2000 Census are about 2 percent *lower* than the weekly earnings of observationally equivalent workers who did not move during that period and lived in the same location in 2000. This evidence is consistent with the view that migrating workers are on average less productive than non-migrating ones in their new location of choice. The model introduced in the next section is consistent with this evidence.

# 3 Model

The model presented in this section builds on the island-model of the labor market developed by Lucas and Prescott (1974) and on McCall (1970)'s intertemporal search model. The force that drives the dynamics of the local labor market in the model is a persistent labor demand shock. Relatively high labor demand shocks generate temporary increases in local wages that are then followed by net inflows of workers. Simultaneously, idiosyncratic wage 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

The economy is populated by a continuum of measure one of locations (islands). A location is denoted by $j \in [0,1]$. All locations are ex-ante identical. Ex-post, locations differ because they are hit by different labor demand shocks, denoted by $z_{jt}$, where $t$ indexes time.[19]

An agent $i$ begins his economic life in period $t$ in a location of birth $j$ by drawing a wage shock $w_{ijt} = \upsilon \overline{w}_{jt}$. The latter is the product of two components. An idiosyncratic component $\upsilon$, which as in Kennan and Walker (2005) captures the match between the worker and the location, and that remains the same as long as the agent stays in the same location. A location-specific component $\overline{w}_{jt}$ common to all workers living in location $j$, which changes over time with the local labor market conditions. At the end of each period, each agent faces a constant probability $\delta$ of death. If he survives, he generates $\left(\delta^{-1} - 1\right)$ children, who will begin their economic life in the location of birth.

In detail, the sequence of events in an agent's life, beginning at birth, is as follows:

- An agent $i$ is born in a location $j$ at the end of period $t-1$.
- At the beginning of $t$, the agent draws the idiosyncratic location-match $\upsilon_i$ from a distribution $f(\upsilon)$, whose mean is normalized to one. The individual shock $\upsilon_i$ and the local component of wages $\overline{w}_{jt}$ determine his wage $w_{ijt}$.
- The agent receives utility $w_{ijt} = \upsilon_i \overline{w}_{jt}$.
- With probability $\delta < 1$ the agent survives into the next period.
- If he survives, the agent has to decide whether to stay in location $j$ or move to another location. If he decides to move to a different location (at the end of period $t$), he pays a moving cost $\kappa$.

[19] In this model I abstract from unemployment. The flow data from the Census describe interstate moves at five years intervals, so a model's period will represent five years. Given that most unemployment spells last only a few weeks, this would create problems in numerically implementing the model. Lkhagvasuren (2005) considers a model that explicitly allows for unemployment in order to explain the large observed cross-state differences in unemployment rates.

- At the end of period $t$, the agent reproduces and generates $\left(\delta^{-1}-1\right)$ children who are born in the new location of choice.

- At the beginning of period $t+1$, if the agent had remained in the same location in which he was living in $t$, he receives momentary utility $w_{ijt+1} = \upsilon_i \overline{w}_{jt+1}$. If the agent has moved to a new location, he draws a new idiosyncratic location-match $\upsilon_i$ (see second bullet above).

**Production.** Aggregate output in location $j$ at time $t$, denoted by $a_{jt}$, is produced according to the following Cobb-Douglas production function:

$$a_{jt} = \left(z_{jt} l_{jt}\right)^{\tau}, \ \tau \in (0,1),$$

where $l_{jt}$ represents labor, measured in efficiency units and $z_{jt}$ is a shock to the productivity of labor located in $j$ at time $t$. The stochastic process $\{z_{jt}\}$ is assumed to be stationary. Additional restrictions on $\{z_{jt}\}$ will be specified later.

The efficiency units of labor $l_{jt}$ located in $j$ at time $t$ are:

$$l_{jt} = y_{jt} \overline{\upsilon}_{jt},$$

where $y_{jt}$ represents the measure of workers located in $j$ at the beginning of $t$ and $\overline{\upsilon}_{jt}$ denotes the efficiency units per worker in that location. Specifically:

$$\overline{\upsilon}_{jt} = \int \upsilon \lambda_{jt}(\upsilon)\, d\upsilon,$$

where $\lambda_{jt}$ denotes the endogenous distribution of idiosyncratic location matches $\upsilon$ in $j$ at the beginning of $t$.[20]

Firms' optimization yields the market wage rate for an efficiency unit of labor:

$$\overline{w}_{jt} = \tau z_{jt}^{\tau} l_{jt}^{\tau-1}. \tag{3}$$

**Individual Mobility.** Let's now consider the problem of an individual agent. In what follows, I concentrate on equilibria in the wage per efficiency unit of labor is equalized across locations:[21]

$$\overline{w}_{jt} = \overline{w}, \ \forall j, t. \tag{4}$$

The value function of an agent characterized by an idiosyncratic match $\upsilon$ with location $j$ is then:

$$V(\upsilon) = \upsilon \overline{w} + \beta\delta \max\left\{V(\upsilon), \int V(\upsilon) f(\upsilon)\, d\upsilon - \kappa\right\}, \tag{5}$$

where the parameter $\beta < 1$ in this equation denotes the discount factor. Notice that, since the wage per efficiency units of labor is the same across locations, the agent must receive the same expected utility in all locations.

The agent will follow a simple cut-off rule in order to determine whether to stay in a location or leave it. Specifically, there exists a cut-off $\widehat{\upsilon}$ such that for $\upsilon < \widehat{\upsilon}$ the agent decides to leave, and for $\upsilon \geq \widehat{\upsilon}$ the agent decides to stay. Given that $\upsilon$ is drawn once and for all in the first period in

[20]Thus, the distribution $\lambda_{jt}$ also includes the workers who have moved to $j$ at the end of $t-1$ and the workers born at the beginning of $t$.

[21]The conditions under which such equilibria exist will be specified later.

which an agent moves to a location (or after he is born), an agent will either leave the location immediately, or stay forever. The cut-off is defined implicitly by the following equation:

$$\widehat{v} = (1-\beta\delta)\left(1-\frac{\kappa}{\overline{w}}\right) + \beta\delta \int \max\{v,\widehat{v}\} f(v)\, dv. \tag{6}$$

The equation above expresses the worker's indifference between staying in the location or moving to a different one.[22] A few observations are in order. First, there exists a value of $\overline{w}$ given by:

$$\overline{w} = \kappa(1-\beta\delta),$$

such that $\widehat{v} = 0$ and for all levels of $\overline{w}$ below this value, the cutoff $\widehat{v}$ is negative. This means that the worker will never leave a location, no matter how low his match with the location. Second, given $\overline{w} > \kappa(1-\beta\delta)$, there exists a unique cut-off $\widehat{v} > 0$. To see this, notice that, for $\overline{w} > \kappa(1-\beta\delta)$, the right-hand side of (6) is strictly positive at $\widehat{v} = 0$ and its derivative with respect to $\widehat{v}$ is smaller than one.

**Equilibrium.** The model is closed by the following equilibrium condition, which pins down the equilibrium value of $\overline{w}$:

$$\int_0^1 y_{jt} dj = 1. \tag{7}$$

According to the latter, the sum of population across locations must be equal to the aggregate population. The latter is constant over time and normalized to one.

### 3.2 Workers' Flows

In this section I describe the dynamics of workers' flows.

**Workers' Flows.** Population in a location evolves according to the law of motion[23]

$$y_{jt} = y_{jt-1} + \frac{1}{\delta}(x_{jt-1} - o_{jt-1}), \tag{8}$$

where $x_{jt}$ and $o_{jt}$ denote, respectively, gross inflows and gross outflows. Net inflows are defined as the difference between gross inflows and gross outflows. Equation (8) implies that the net inflow to a location $j$ between periods $t-1$ and $t$ is given by:

$$d_{jt-1} = \delta(y_{jt} - y_{jt-1}). \tag{9}$$

In each period $t$, there are two categories of workers who *might* leave a location $j$: workers born in $j$ at the end of $t-1$ and immigrants who arrived in $j$ at the end of $t-1$. Let the size of the former group be denoted by $n_{jt-1}$, where:

$$n_{jt-1} = \left(\frac{1}{\delta} - 1\right)(\delta y_{jt-1} + x_{jt-1} - o_{jt-1}). \tag{10}$$

[22] Equation (6) makes clear that the heterogeneity in ability can be reinterpreted as heterogeneity in moving costs.
[23] Notice that the change in population in a location between $t-1$ and $t$ can be decomposed as follows:

$$y_{jt} - y_{jt-1} = \underbrace{-(1-\delta)\, y_{jt-1}}_{\text{deaths}} \underbrace{- o_{jt-1}}_{\text{outflows}} + \underbrace{\left(\frac{1}{\delta}-1\right)(\delta y_{jt-1} - o_{jt-1} + x_{jt-1})}_{\text{births}} + \underbrace{x_{jt-1}}_{\text{inflows}}.$$

The size of the latter group is simply $x_{jt-1}$. Gross outflows from location $j$ between $t$ and $t+1$ are then:

$$o_{jt} = \delta \left( n_{jt-1} + x_{jt-1} \right) F \left( \widehat{\upsilon} \right). \tag{11}$$

Notice that $F\left(\widehat{\upsilon}\right)$ in (11) represents the probability that an agent draws a shock below the migration cutoff $\widehat{\upsilon}$. This probability is constant over time.

**Efficiency Units.** In order to complete the characterization of the model's flows we need to determine the dynamics of population $y_{jt}$ in a location. Notice that the equality among marginal products (equation 4) yields the following expression for the efficiency units of labor located in $j$ at time $t$:

$$y_{jt}\overline{\upsilon}_{jt} = \left( \frac{\tau z_{jt}^{\tau}}{\overline{w}} \right)^{\frac{1}{1-\tau}}. \tag{12}$$

The aggregate efficiency units of labor in location $j$ at the beginning of time $t$ satisfy the following law of motion:

$$y_{jt}\overline{\upsilon}_{jt} = \delta y_{jt-1}\overline{\upsilon}_{jt-1} + n_{jt-1} + x_{jt-1} - o_{jt-1}\frac{1}{F(\widehat{\upsilon})}\int_{0}^{\widehat{\upsilon}} \upsilon f(\upsilon) d\upsilon \tag{13}$$

The left-hand side of equation (13) represents the aggregate efficiency units of labor at the beginning of $t$. The latter equals the aggregate efficiency units of the surviving workers located in $j$ in $t-1$ *plus* the efficiency units of new born agents and incoming workers *minus* the efficiency units of workers that leave the location at the end of period $t-1$.[24]

### 3.3 Equilibrium

The equilibrium of the model can be found analytically. To do so, first use equation (12) to replace $y_{jt}\overline{\upsilon}_{jt}$ and its lagged value in equation (13). Then, replace in (13) the definition of $n_{jt-1}$ (equation 10) and $x_{jt-1}$ from equation (8). This yields:

$$y_{jt} = \delta y_{jt-1} - \varphi o_{jt-1} + \left( \frac{\tau}{\overline{w}} \right)^{\frac{1}{1-\tau}} \left( z_{jt}^{\frac{\tau}{1-\tau}} - \delta z_{jt-1}^{\frac{\tau}{1-\tau}} \right), \tag{14}$$

where the coefficient $\varphi$ is defined as:

$$\varphi \equiv 1 - \frac{1}{F(\widehat{\upsilon})}\int_{0}^{\widehat{\upsilon}} \upsilon f(\upsilon) d\upsilon.$$

Similarly, replace $n_{jt-1}$ and $x_{jt-1}$ from equations (10) and (8), respectively, into equation (11), to obtain:

$$o_{jt} = \delta \left( o_{jt-1} + y_{jt} - \delta y_{jt-1} \right) F\left( \widehat{\upsilon} \right). \tag{15}$$

Equations (14) and (15), together with the equilibrium condition (7) and the stochastic process for $z_{jt}$, completely characterize the equilibrium of the model. In particular, to solve for $\overline{w}$, it is enough to integrate the left and right-hand side of equations (14) and (15) over the distribution of locations, imposing the equilibrium condition (7):

$$\overline{w} = \tau \left[ \frac{1 - \delta F\left(\widehat{\upsilon}\right)}{1 - \delta\left(1 - \varphi\right) F(\widehat{\upsilon})} \int_{0}^{1} z_{jt}^{\frac{\tau}{1-\tau}} dj \right]^{1-\tau}.$$

[24] Notice that the workers who draw the idiosyncratic shock $\upsilon$ in location $j$ for the first time in period $t$ ($n_{jt-1}+x_{jt-1}$) draw from the unconditional distribution $f(\upsilon)$. Thus, their average efficiency units is one.

It is easy to show that equations (14) and (15) describe a stable system of difference equations.

So far, I have not imposed restrictions on the process for $z_{jt}$, except for stationarity. However, in order for the equilibrium just described to be an equilibrium, it is necessary that some additional conditions be satisfied. In particular, gross inflows and gross outflows must always be non-negative. The appendix show that a set of *sufficient* conditions for this is that:

- The stochastic process $\{z_{jt}\}$ and the model's parameters are such that it is always the case that:
$$\Delta_{\min} < z_{jt}^{\frac{\tau}{1-\tau}} - \delta z_{jt-1}^{\frac{\tau}{1-\tau}} < \Delta_{\max},$$
for some numbers $\Delta_{\min} > 0$ and $\Delta_{\max} > 0$.

- The parameter $\delta$ is "small enough". The appendix gives the exact restrictions that $\delta$ has to satisfy.

# 4 Discussion

Before proceeding it is worth discussing some of the modelling choices that I have made.

**Demand vs. Supply Shocks.** The first modelling choice is that the driving force of the model is represented by local labor *demand* shocks, as opposed to local labor *supply* shocks. In doing so, I am not considering the possibility that the high net inflows of population experienced by states such as Nevada and Arizona in recent decades, might be driven by local amenities (e.g. warm winters). It turns out that this choice, while being empirically supported, is not restrictive. It is in fact possible to add amenities to the model of section (3) and show that it gives rise to identical implications for gross and net flows as the model driven exclusively by demand shocks. To see this, let a location $j$ be characterized at time $t$ by a vector of amenities $\mathbf{k}_{jt}$ which can vary stochastically over time, and that is observed at the end of $t-1$ by prospective movers.[25] I assume that amenities affect agents' utility in a multiplicative fashion. An agent with match shock $\upsilon$ that lives in a location where the wage per efficiency unit of labor is $\overline{w}_{jt}$ has instantaneous utility $\upsilon \overline{w}_{jt} h(\mathbf{k}_{jt})$. As in section (3.1), I focus on equilibria in which $\overline{w}_{jt} h(\mathbf{k}_{jt})$ is equalized across locations in every period:

$$\overline{w}_{jt} h(\mathbf{k}_{jt}) = \overline{w}. \tag{16}$$

Equation (16) captures the compensating effect of amenities on wages: locations with better amenities (i.e., higher $h(\mathbf{k}_{jt})$) pay lower wages. The rest of the analysis is identical to before. In particular, using the definition of $\overline{w}_{jt}$, one obtains that the aggregate efficiency units located in $j$ are now:

$$y_{jt}\overline{\upsilon}_{jt} = \left( \frac{\tau z_{jt}^{\tau} h(\mathbf{k}_{jt})}{\overline{w}} \right)^{\frac{1}{1-\tau}}, \tag{17}$$

so that variations in amenities can also play the role of driving force behind population flows, in addition to labor demand shocks. In particular, the dynamics of net and gross flows in the model driven exclusively by amenity shocks become observationally equivalent to the ones generated by the model of section (3) if $\mathbf{k}_{jt}$ is a scalar and $h(\mathbf{k}_{jt}) = \mathbf{k}_{jt}^{\tau}$.

The upshot of this discussion is that the model with amenities produces qualitatively identical implications for workers' flows as the model driven by labor demand shocks. In order to solve this identification problem, it is necessary to consider the relationship between real earnings and net

[25] For example, the introduction of air conditioning has significantly improved living conditions in the South-West of the U.S. during summer months. This can be interpreted as a change in $\mathbf{k}$.

flows. If the driving force of workers' flows were amenity shocks, one would expect a negative cross-sectional correlation between real earnings and net flows of workers. The sign of this correlation is implied by equations (16) and (17).

The available evidence seems, *prima facie*, to suggest against this hypothesis. Table 7 reports the cross-sectional correlation coefficients among the following state-level variables: gross inflows, gross outflows, net flows, average (log) real weekly earnings (these four variables are from the 2000 Census), annual heating and cooling-degree days.[26]

Table 7
Cross-Sectional Correlations (Census 2000)

| | Outflow Rate | Inflow Rate | Net Flow Rate | Average Log Wages | Cooling Degree Days | Heating Degree Days |
|---|---|---|---|---|---|---|
| Outflow Rate | 1 | 0.66*** | 0.25* | −0.17 | 0.25* | −0.31** |
| Inflow Rate | | 1 | 0.89*** | 0.25* | 0.46*** | −0.47*** |
| Net Flow Rate | | | 1 | 0.43*** | 0.44*** | −0.42*** |
| Average Log Wages | | | | 1 | 0.37*** | −0.25* |
| Cooling Degree Days | | | | | 1 | −0.81*** |
| Heating Degree Days | | | | | | 1 |

*** significant at 1% level, ** significant at 5% level, * significant at 10% level.

The Table clearly show how states with larger positive net flows of workers tend to be characterized by relatively higher real earnings (Figure 4).[27] These are also the states with better amenities, as measured by the number of cooling and heating degree-days (Figure 5). These correlations are, *prima facie*, inconsistent with theories that postulate the existence of compensating differentials in earnings across locations (see e.g., Roback (1982)).

Of course, the question of why certain states experience large positive and persistent labor demand shocks remains open. No attempt is made to answer this difficult question in this paper. The approach taken here will be to first use these net flows to back out the underlying local labor demand shocks. These shocks will then be fed into the model and the latter will be evaluated on the basis of its cross-sectional implications for the gross flows of workers.

**Persistent Differences Across States.** A crucial aspect of the modelling strategy concerns the nature of heterogeneity across states. The model above describes a *stationary* economy in which all locations are ex-ante *identical*. Differences across locations are persistent but not permanent. In

[26] I have used the dummy coefficient $\mu_j$ as a measure of average (log) weekly earnings in state $j$ in 1999. The annual number of cooling and heating degree-days are from the U.S. Historical Climatography Series 5-2 and 5-1. They are averages over the period 1931-2000. Given that there might be considerable within-state variation in temperature, the series are constructed using a population-weighted aggregate temperature for each state, with weights given by the Census 2000 population. The number of annual cooling and heating degree-days for a state $j$ in year $t$ are defined as:

$$\text{cooling degree days} = \sum_{d=1}^{365} \max\{x_{jtd} - 65, 0\},$$
$$\text{heating degree days} = \sum_{d=1}^{365} \max\{65 - x_{jtd}, 0\},$$

where $x_{jtd}$ is the average daily temperature in day $d$ of year $t$ in state $j$.

[27] A similar conclusion has also been reached by Topel (1986) and Blanchard and Katz (1992), who find it more support for local shocks to labor demand as the driving force behind net flows of workers across states.

particular, a location's population share and earnings per capita tend to return over time to their long-run value, which is common across locations. This modelling choice, although convenient for many reasons, deserves some further comment.

Consider first the evidence concerning per-capita income or wages. Barro and Sala-i-Martin (1992) and (1991), using state-level data dating back to 1840, have shown how states with lower initial incomes per capita have subsequently grown faster towards a common income level than initially richer states.[28] Blanchard and Katz (1992) also find strong evidence of convergence of manufacturing wages across U.S. states in the post-WWII period. Differently from Barro and Sala-Martin, though, in their empirical and theoretical analysis, Blanchard and Katz specify states' relative wages as stationary processes around state-specific means.[29]

Blanchard and Katz (1992, page 5) also document that, for the period 1950-1990, "U.S. states have experienced large and sustained differences in employment growth rates." They capture this observation in a formal way by modelling the growth rates of states' employment shares as stationary processes with state-specific means. Of course, since a state's employment share is bounded from above by one this formalization cannot be taken literally, but just as a convenient way of capturing the persistent differences in employment growth across states. This specification is not problematic for Blanchard and Katz because, in their reduced form model, they never impose the condition that the employment shares must be smaller than one. In the model above, instead, employment shares, rather than their first differences, are assumed to be stationary. This is a necessary assumption for a well-defined equilibrium. Of course, it is always possible to set the parameters controlling the degree of mean reversion to a small enough value, so that in practice it might be impossible to distinguish one specification from the other. In summary, both with respect to relative wages and employment shares, I interpret observed differences across states as stemming from slow transitions, rather than from time-invariant features.

**Firms' Mobility vs. Workers Mobility.** The model of the previous section abstracts from capital mobility. This is without loss of generality. It is easy to introduce capital in the production function. Capital mobility would then imply the equalization of the marginal product of capital across locations. This marginal product condition can then be used to solve for the amount of capital in each location as a function of its efficiency units of labor. The location's capital stock can be replaced in the original production function, leading to a new production function that depends only on labor. The model is also consistent with the alternatively assumption that capital is fixed in each location.[30]

**The Labor Market.** In the model of the previous section, two types of agents are active in the labor market at a point in time. The first one is young individuals that are "just born", and draw

[28]To be more precise, Barro and Sala-i-Martin's evidence points to what is usually called in the growth literature "$\beta$-convergence". The latter denotes the tendency for states with lower incomes per capita to grow faster than states in which income per capita is relatively large.

[29]It is worth noticing that both Barro and Sala-i-Martin and Blanchard and Katz cannot control for differences in price levels across states.

[30]Lee (2004) provides comprehensive evidence regarding plant relocation in the manufacturing sector for the period 1972-1992. He finds (page 17) that, in a five-year period, "plant relocations account for about 7 percent of variations in net employment growth across states. The remaining 93 percent is accounted for by within-state changes such as employment growth within continuing plants, de novo plant openings, permanent closings without relocation, and intrastate plant relocation." This evidence suggests that movements of existing plants play a relatively minor role in explaining the differential growth of states' employment. It does not rule out, though, that the opening of new plants could play a much more significant role. Blanchard and Katz (1992) provide evidence that suggests that neither plant relocation nor new plant openings play an important role in states' adjustment to local labor demand shocks. They estimate local labor demand equations and find that local employment reacts strongly to local labor demand shocks, while job creation responds weakly to movements in local wages.

their first wage shock. The second is movers from other locations who draw new wage shocks. If the idiosyncratic shock a worker receives is high enough, the worker will stay in the location as long as he lives. Otherwise, he will move to a different location, draw a different shock, etc. An important assumption here is that there is no search at the local level, i.e., workers sample from the local wage offer distribution only once. This is a way to capture the fact that there is a local component to wages. For example, a company's manager might realize that he cannot advance in his career in the current location, but can move within the company to a different location, be promoted and given a salary raise. In the model the worker has to move to find out what his wage will be in the new location. While this is unlikely to be true in most cases, the worker is moving to a new job and there is going to be uncertainty regarding the job's features (colleagues, working hours, possibility of further advancements, etc.). This uncertainty generates uncertainty in wages, as the latter might reveal ex-post to be too low, given the job's characteristics.

# 5 The Model and the Data

In this section I ask whether the model is qualitatively consistent with the stylized facts described in section (2). Define the net inflow rate $\widehat{d}_{jt}$ as:[31]

$$\widehat{d}_{jt} \equiv \frac{x_{jt} - o_{jt}}{\delta y_{jt}}.$$

Using equations (14) and (8) one obtains the following expression for the net inflow rate as function of the state variables $\widehat{o}_{jt}$ and $y_{jt}$:

$$\widehat{d}_{jt} = \delta - 1 - \varphi \delta \widehat{o}_{jt} + \frac{1}{y_{jt}} \left( \frac{\tau}{\overline{w}} \right)^{\frac{1}{1-\tau}} \left( z_{jt+1}^{\frac{\tau}{1-\tau}} - \delta z_{jt}^{\frac{\tau}{1-\tau}} \right), \tag{18}$$

where $\widehat{o}_{jt}$ denotes the gross outflow rate. Of course, the gross inflow rate is just:[32]

$$\widehat{x}_{jt} = \widehat{o}_{jt} + \widehat{d}_{jt}.$$

From equations (15) and (14), the outflow rate from location $j$ and $j$'s population jointly evolve according to:

$$\begin{aligned} y_{jt+1} &= \delta y_{jt} \left(1 - \varphi \widehat{o}_{jt}\right) + \left( \frac{\tau}{\overline{w}} \right)^{\frac{1}{1-\tau}} \left( z_{jt+1}^{\frac{\tau}{1-\tau}} - \delta z_{jt}^{\frac{\tau}{1-\tau}} \right), \\ \widehat{o}_{jt} &= F\left(\widehat{v}\right) - \delta F\left(\widehat{v}\right) \frac{y_{jt-1}}{y_{jt}} \left(1 - \widehat{o}_{jt-1}\right). \end{aligned} \tag{19}$$

[31] In order to make the definition of net inflow rate consistent with the one employed with the Census data, I define the net inflow rate in this way rather than as

$$\frac{x_{jt+1} - o_{jt}}{y_{jt}}.$$

The difference is due to the fact that the Census records a move relative to an individual's location of residence 5 years before. Given that the Census is available only every 10 years, I have standardized the observed net flows by the number of people living in location $j$ at the beginning of $t-1$ who were also alive in period $t$. That is, all the flows have been standardized by $\delta y_{jt-1}$, rather than by $y_{jt-1}$. The latter variable is not observable in the data, but the former is.

[32] Formally, their definitions are:

$$\widehat{x}_{jt} = \frac{x_{jt}}{\delta y_{jt}},\ \widehat{o}_{jt} = \frac{o_{jt}}{\delta y_{jt}}.$$

The best way to appreciate the dynamics of the model is to consider its response to a on-time shock. Even of this can be done analytically, it is simpler to consider a parameterized version of the model and consider its properties in this context. The qualitative properties of the model do not depend on the specific parameters under consideration.

**Parameterization.** The following parameters have been selected:

- A period in the model is taken to represent 5 years. An individual's working life in the data I consider is about 30 years, or 6 model-periods. Thus, I set the constant probability of survival $\delta$ equal to 0.83, so that the average lifetime for an individual in the model is 5 periods.
- The labor share parameter $\tau$ is set equal to 0.5.
- Given $\delta$, the hazard rate $F(\widehat{v})$ is set to match, on average, the observed interstate migration rate in the U.S. economy, in the 1995-2000 period:

$$\frac{(1-\delta)\,F(\widehat{v})}{1-\delta F(\widehat{v})} = 0.086 \rightarrow F(\widehat{v}) = 0.36.$$

- Assume that the distribution $f(v)$ is lognormal with mean normalized to 1 and variance $\exp\left\{\sigma_v^2\right\} - 1$. The parameter $\sigma_v$ determines the dispersion in the idiosyncratic component of wages within a location.

  It is identified by the fraction of observed inequality in residual earnings that can be attributed to location effects, after controlling for differences in mean wages across states and workers' observable characteristics. This figure is reported by Kennan and Walker (2005). Matching this figure yields $\sigma_v = 0.32$. Notice that, given $\sigma_v$ and $F(\widehat{v})$, it is possible to recover $\widehat{v}$, and use it to compute $\overline{w}$.
- The shock process $\{z_{jt}\}$ is assumed to be such that

$$z_{jt+1} - \delta z_{jt} = (1-\delta)\varepsilon_{jt+1}, \qquad (20)$$

  where

$$\varepsilon_{jt+1} = \rho\varepsilon_{jt} + u_{jt+1}.$$

  The unconditional mean of $\varepsilon_{jt+1}$ is normalized to one. It's important to notice that if $\rho = 0$ net flows in the model would be negatively correlated over time. This is due to the decreasing marginal product of labor. A positive productivity shock this period leads to a net inflow of workers until the marginal product of labor is equalized with the rest of the economy. In the following period, though, productivity $z$ is below steady state due to the fact that $\delta < 1$ in equation (20). Thus, being able to account for the positive autocorrelation of net flows over time requires the shock $\varepsilon_{jt+1}$ to be persistent over time. The parameter $\rho$ is set to 0.95.

**The Mechanics of Workers' Flows.** Figure 6 plots the period-by-period response of outflow, inflow and net flow rates (represented as deviations from their steady state values) to a one-time unanticipated shock $u_{j1} > 1 - \rho$. In detail:

- In period 1 gross outflows stay constant at their steady state level, while at the end of period 1, there is a net inflow of workers who are attracted by the expected positive conditions of the local labor market for period 2.[33] Thus, gross inflows increase by exactly the same amount as net inflows.

[33] Remember that the timing of the model is such that workers cannot, by assumption, move in from other locations to take advantage of the positive labor market conditions in period 1.

- In period 2, the shock $u_{j2}$ is back to 1. Net flow rates, however, are driven by the expected growth in $z_{jt}$. Given that the growth rate shock $\varepsilon_{jt+1}$ is positively autocorrelated, net flows in period 2 are also above their steady state value. Gross outflows tend to rise above steady state due to the relatively large gross inflow of workers in the previous period. Some of these workers, in fact, are ex-post unlucky and decide to migrate again. Notice that the incoming workers leave the location at the same rate as individuals that are just born in the location. What makes the outflow rate increase is that the average outflow rate for the economy as a whole is a weighted average of the outflow rate of incoming and new born agents ($\delta F(\widehat{v})$) and incumbent agents. Since the incumbents' outflow rate is zero, the average outflow rate increases after an increase in gross inflows. As a result, the gross inflow rate in period 2 exceeds the net inflow rate.

- The following periods are similar to period 2, with net flows moving back toward steady state, gross flows exceeding steady state, and gross inflows exceeding net inflows.

What moments does this adjustment path imply? Compare first the standard deviation of net inflow, gross inflow and gross outflow rates. On impact, following the positive labor demand shock, net inflows rise while gross outflows remain constant. Over time gross outflows respond to the higher inflow of workers. The fact that gross outflows are above steady state in the periods following the shock implies that gross inflows exceed net inflows. Thus, gross inflows are more volatile than net inflows, and the latter are more volatile than gross outflows.

In terms of correlations, in the period following the shock all three flows tend to move together leading to a high correlation between any two of them. In the data only gross and net inflows are highly correlated. Gross inflows and outflows are positively correlated, but less than the amount predicted by the model. The latter predicts that net flows and gross outflows are positively correlated in the cross-section, while they tend to be slightly negatively correlated in the data. Finally, gross and net flows are highly positively correlated over time in the model and in the data.

**Average Wages.** Figure 7 represents the response of average log wages in a location. The latter are given by:

$$\ln \overline{w} + \int_0^\infty \ln(v)\,\lambda_{jt}(v)\,dv. \tag{21}$$

At each point in time, locations will be characterized by different average wages because the distribution of idiosyncratic efficiency shocks, $\lambda_{jt}(v)$, will differ across locations, due to their different histories of $u$-shocks. The $u_{j1}$ shock induces a gross inflow of workers in the location. These workers draw shocks from the unconditional distribution $F(v)$, which, due to the selection associated with workers' moving decisions, has a lower mean than the distribution of $v$ across incumbent workers. To see this, notice that, in steady state, the latter is:

$$\int_0^\infty \ln(v)\,\lambda_j(v)\,dv = -\frac{\sigma_v^2}{2} + \frac{\delta F(\widehat{v})}{1-\delta F(\widehat{v})}\left\{-\frac{\sigma_v^2}{2} - \frac{1}{F(\widehat{v})}\int_0^{\widehat{v}} \ln v f(v)dv\right\}, \tag{22}$$

while the mean of the $F$ distribution is $-\sigma_v^2/2$. The term in curly brackets in (22) is positive due to the selection effect mentioned above. It follows that the incoming workers will have, on average, lower wages than incumbents. This is consistent with the evidence of section (2). Notice that, while the cross-sectional data cannot provide more information on the nature of this negative selection, the simple model of migration presented here helps ruling out the alternative interpretation of this evidence. That is, if workers were ex-ante different, so that their productivity in the source location were positively correlated with their productivity in the location of destination, the model would predict that low ability workers would have a lower tendency to move than high ability ones.[34] Thus, if the pool of migrating workers was mainly populated by high ability types, we should observe a wage premium associated with migration, as opposed to a wage discount as found in the data.

## 6 Discussion and Extensions

[TO BE COMPLETED]

There are two features of the data that the model cannot reproduce. First, the model predicts that outflow rates in a location are highly positively correlated with both gross and net inflows. This is at odds with the data, as observed outflow rates display a small positive correlation with gross inflows and a small negative correlation with net inflows. In order to address this problem, it might be useful to introduce in the model a shock that affects outflows directly. For example, in the current version of the paper, if an agent draws an idiosyncratic shock $\upsilon > \widehat{\upsilon}$, he chooses to stay in the location where the shock is drawn until he dies. Alternatively, one could assume that the duration of all idiosyncratic matches in a location is stochastic. In this setting, locations hit by aggregate "destruction" shocks would experience a gross outflow of workers, which is likely to reduce the correlation between gross outflows and the other flows, in addition to making outflows more volatile. Future versions of the paper will contain this extension.

Second, the model cannot reproduce the observed dispersion in average real earnings across states and the observed correlation between wages and net flows. The observed dispersion of average earnings can be attributed to two potentially complementary causes. On the one hand, population flows might be subject to adjustment costs that slow down the process of $\beta$-convergence of earnings across locations. This will result into larger cross-sectional dispersion at a point in time. On the other hand, differences in real earnings might capture differences in amenities across locations. As discussed above, it is not easy to determine a-priori which amenities people value. The positive correlation between real earnings and measures of "good weather" across states suggests that weather might not be the appropriate amenity to consider. Of course, it could be that the observed differences in real earnings are due to measurement error because of the difficulty of accurately measuring the level of prices in different locations. The availability of better data is crucial to be able to make progress in this area. In order to dig deeper into this issue, I am in the process of collecting local wage deflators for 1989, 1979 and 1969. Unfortunately, these deflators, computed by ACCRA, are not comparable through time for a given location, but only across locations at a given point in time. Despite this limitation, they can be used to compute rank correlations of average state wages across decades, to get a measure of how persistent these wage differences are.

## 7 Summary

This paper makes two contributions. First, it presents the main stylized facts about net and gross flows of workers across U.S. states. Then, it presents a simple model of workers' flows across locations. Preliminary quantitative results are encouraging as the model seems able to reproduce the main features of the data.

[34]This is because high ability is equivalent to a lower mobility cost $\kappa$.

**References**


Barro Robert and Xavier Sala-i-Martin (1991): "Convergence Across States and Regions," *Brooking Papers on Economic Activity* 1991, 107-158.

Barro Robert and Xavier Sala-i-Martin (1992): "Convergence," *Journal of Political Economy* 100, 223-251.

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

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

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

Davis, Steven and John Haltiwanger (1992): "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *Quarterly Journal of Economics* 107, 819-863.

Duffie, Darrell and Kenneth Singleton (1993): "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica* 61, 929-952.

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.

Hassler, John, Rodriguez-Mora, Jose', Storesletten, Kjetil, and Fabrizio Zilibotti (2005): "A Positive Theory of Geographic Mobility and Social Insurance," *International Economic Review 46*, 263-303.

Kennan, John and James Walker (2005): "The Effect of Expected Income on Individual Migration Decisions, " Working Paper, University of Wisconsin-Madison.

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

Lee Bong-Soo and Beth Ingram (1991): "Simulation Estimation of Time-Series Models," *Journal of Econometrics* 47, 197-205.

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

McCall, John (1970): "Economics of Information and Job Search," *Quarterly Journal of Economics* 84, 113-126.

OECD (2000): *Employment Outlook*, Paris.

Roback, Jennifer (1982): "Wages, Rents and the Quality of Life," *Journal of Political Economy* 90, 1257-1278.

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.

# A Data Appendix

### Sample Selection

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:

- were between 27 and 60 years of age (as of their last birthday);
- were not living in group quarters;
- were in the labor force but not in the armed forces;
- if foreign-born, had immigrated to the U.S. at least 5 years before the Census year;
- were not living abroad 5 years before the Census year;
- were not living in the Census year or 5 years before the Census year, in either Alaska, Hawaii, or the District of Columbia.

In what follows, I will refer to the selected sample as the "population". The number of selected individual observations is , representing  million people.

### Measures of Flows

In order to construct measures of gross and net flows I adopt the following procedure. Individual $i$ is observed living in state $j$ in Census year $\tau$. The same individual is also observed living in state $k$ in year $\tau - 5$. Construct an indicator function $I_{i\tau}(j)$ for each individual $i$ such that $I_{i\tau}(j) = 1$ if individual $i$ was recorded as living in location $j$ in Census year $\tau$ and zero otherwise. Also, define an indicator function $\overline{I}_{i\tau}(j)$ such that $\overline{I}_{i\tau}(j) = 1$ if individual $i$, interviewed in Census year $\tau$, reported living in location $j$ in year $\tau - 5$. Total outflow of population from location $j$ between $\tau - 5$ and $\tau$ is then defined as

$$out_{j\tau} = \sum_{k \neq j} \sum_{i} \mu_{i\tau} I_{i\tau}(k) \overline{I}_{i\tau}(j),$$

where $\mu_{i\tau}$ is the person weight (`perwt`) assigned by the year $\tau$ Census to individual $i$. The total inflow of population into location $j$ between $\tau - 5$ and $\tau$ is analogously defined as:

$$in_{j\tau} = \sum_{k \neq j} \sum_{i} \mu_{i\tau} I_{i\tau}(j) \overline{I}_{i\tau}(k).$$

Let $y_{j\tau}$ denote the population living in location $j$ in Census year $\tau$:

$$y_{j\tau} = \sum_{i} \mu_{i\tau} I_{i\tau}(j).$$

I also denote by $\overline{y}_{j\tau}$ the total population that was interviewed in year $\tau'$s Census that was living in location $j$ in year $\tau - 5$:

$$\overline{y}_{j\tau} = \sum_{i} \mu_{i\tau} \overline{I}_{i\tau}(j).$$

An outflow *rate* from location $j$ between $\tau - 5$ and $\tau$ is then defined as follows:

$$\widehat{o}_{j\tau} = \frac{out_{j\tau}}{\overline{y}_{j\tau}}.$$

Analogously an inflow *rate* into location $j$ between $\tau - 5$ and $\tau$ is defined as

$$\widehat{x}_{j\tau} = \frac{in_{j\tau}}{\overline{y}_{j\tau}}.$$

The net flow *rate* into location $j$ between $\tau - 5$ and $\tau$ is defined as the difference between inflow and outflow rates:

$$\widehat{d}_{j\tau} = \widehat{x}_{j\tau} - \widehat{o}_{j\tau}.$$

**Accounting for Heterogeneity**

There are (at least) two sources of heterogeneity I need to worry about. The first concerns heterogeneity among locations (U.S. states) in the demographic composition of their population. For example, if state A has a younger workforce than state B, and, if younger individuals are characterized by higher mobility, then measures of outflows from state A will be higher than from state B due to these demographic differences. Given that, in the model, all individuals have the same age I need to control for these differences when constructing measures of outflows and inflows. The second, equally important, issue is that while in the model all locations are identical in terms of their "physical attributes", in the data, instead, different states are characterized by very different geographic features, sizes, etc. For example, outflow rates might be higher from smaller (size being measured in terms of land area) states.
In practice, adjusting the data to take into account these concerns is not straightforward, because while some of the heterogeneity across states can be safely taken as exogenous (e.g., their land area), other features such as their demographic composition are likely not to be so.
In what follows I proceed in two steps. First, in order to construct measures of population flows that are free from potential composition effects, I divide the population in each location (state) in different cells, defined by the following characteristics in the Census year:

- age (`age`); 7 age groups: 27-31, 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 (`ind`), including the unemployment state; 14 industries: (1) unemployed, (2) agriculture, fishing, forestry and mining, (3) utilities, (4) construction, (5) manufacturing, (6) wholesale and retail sales, (7) transportation and wharehousing, (8) information and communication, (9) finance, insurance, real estate and leasing, (10) professional, scientific, management, (11) educational, health, social service, (12) arts, entertainment, recreation, (13) other services, (14) public administration.

Denote each cell by $g$ and the collection of cells by $G$. There are 490 cells. For each cell $g$ it is possible to construct the equivalents of total outflows and inflows defined above in the following way:

$$\begin{aligned} out_{jg\tau} &= \sum_{k\neq j}\sum_{i\in g}\mu_{i\tau}I_{i\tau}(k)\,\overline{I}_{i\tau}(j)\,, \\ in_{jg\tau} &= \sum_{k\neq j}\sum_{i\in g}\mu_{i\tau}I_{i\tau}(j)\,\overline{I}_{i\tau}(k)\,, \\ \overline{y}_{jg\tau} &= \sum_{i\in g}\mu_{i\tau}\overline{I}_{i\tau}(j)\,. \end{aligned}$$

Group-specific outflow and inflow rates are then defined as

$$\widehat{o}_{jg\tau} = \frac{out_{jg\tau}}{\overline{y}_{jg\tau}},\ \widehat{x}_{jg\tau} = \frac{in_{jg\tau}}{\overline{y}_{jg\tau}}.$$

Define the population share of cell $g \in G$ over the U.S. population:

$$\upsilon_{g\tau} = \frac{\sum_j \overline{y}_{jg\tau}}{\sum_j \sum_g \overline{y}_{jg\tau}}.$$

When aggregating the inflow and outflow measures across cells, I use the weight $\upsilon_{g\tau}$ to control for composition effects. So, the adjusted outflow and inflow rates for location $j$ are defined as:

$$\begin{aligned} \widehat{o}^{adj}_{j\tau} &= \sum_g \upsilon_{g\tau} \times \widehat{o}_{jg\tau}, \\ \widehat{x}^{adj}_{j\tau} &= \sum_g \upsilon_{g\tau} \times \widehat{x}_{jg\tau}. \end{aligned}$$

Similarly, it is possible to define net flows.
The second step of the procedure consists of controlling for geographical and historical differences among U.S. states and the effect these might have on population flows. I do this by running separate cross-sectional regressions of inflow and outflow rates on the following state-level variables: (1) land area, (2) year when state joined the U.S., (3) number of metropolitan areas within a state with population larger than 500,000 (computed including all inhabitants) in the 2000 Census. The inflow and outflow rates presented in the main body of the paper are the residuals from these regressions.

**Real Weekly Earnings**

Data on workers' weekly earnings were computed from the Census 2000 by summing annual wage income (`incwage`), business and farm income (`incbus00`), and welfare income (`incwelfr`), and dividing the sum by the number of weeks worked (`wkswork1`). Each source of income refers to the year 1999. 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 for 1999. In a few instances reported earnings by self-employed individuals were negative, and these observations have been dropped. Given that earnings refer to 1999 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 1999. I have also dropped these individuals from the sample.

The earnings data were deflated using the ACCRA Cost of Living index for the third quarter of 1999. This index measures relative price levels (gross of taxes) for consumer goods and services (including housing) in a number of U.S. cities.[35] This number varies from quarter to quarter, and the third quarter of 1999 was selected to maximize coverage of locations (330). Using information on the workers' metropolitan statistical area or PUMA of residence in 2000 (found in the Census), I have matched workers with a value of the Cost of Living index in their area of residence. Unfortunately, the limitations in the coverage of the ACCRA index prevent one from being able to do so for all workers in the Census sample. In particular, only 53 percent of the workers could be matched with a value of the Cost of Living Index. The results reported in the main text of the paper refer only to these workers (about 2.5 million individuals).

The logarithm of real weekly earnings was regressed on the following variables: 48 dummies for workers' state of residence in 2000 (`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", from `raced`), 14 sectoral dummies (from `ind`), and 26 occupational dummies (from `occ`). The $R^2$ of this regression was 30 percent.

**March Current Population Survey**

March CPS data for the years 1999-2003 were used to compute the fraction of individuals moving across U.S. states for job-related reasons. During these years the March CPS contains a question regarding an individual's primary reason for changing residence with respect to the previous year. The Census does not contain such question. The March CPS questionnaire identifies 16 different primary reasons for moving, with "New job or job transfer" (33 percent of the answers) representing the most-frequent single answer, followed by "Other family reasons" (13 percent). In order to compute the fraction of individuals moving for job related reasons I have first applied the sample selection criteria listed above to the March CPS data. Then, the survey's different 16 reasons for moving were aggregated into two categories: job-related and non job-related reasons. The aggregation is relatively straightforward, with the exception of moves motivated by the desire of new, better or cheaper housing. I have included those reasons in the job-related move. The rationale for this choice is that the wage data used to calibrate the model have been deflated using price indices that include housing prices.

# B Existence of Equilibrium

In this section I prove the sufficient condition that appears in section (3.3). Summing gross outflows (equations 15 and 14) and net inflows (equation 8) one obtains an expression for gross inflows. The latter can be manipulated to yield the law of motion:

$$x_{jt} = \delta\left(1-\varphi\delta\right)F\left(\widehat{v}\right)x_{jt-1} - \delta\left(1-\delta\right)\left[1-\left(1-\varphi\delta\right)F\left(\widehat{v}\right)\right]y_{jt} + \delta\left(\frac{\tau}{\overline{w}}\right)^{\frac{1}{1-\tau}}\left(z_{jt+1}^{\frac{\tau}{1-\tau}} - \delta z_{jt}^{\frac{\tau}{1-\tau}}\right).$$

Now, postulate that $x_{jt-1} \geq 0$ and derive sufficient conditions under which $x_{jt} \geq 0$. For this to be the case, it is necessary and sufficient that:

$$\left(\frac{\tau}{\overline{w}}\right)^{\frac{1}{1-\tau}}\left(z_{jt+1}^{\frac{\tau}{1-\tau}} - \delta z_{jt}^{\frac{\tau}{1-\tau}}\right) \geq \left(1-\delta\right)\left[1-\left(1-\varphi\delta\right)F\left(\widehat{v}\right)\right]y_{jt}. \tag{23}$$

Assume that the stochastic process for $\{z_{jt}\}$ is such that there exist $\Delta_{\min} > 0$ and $\Delta_{\max} > 0$ such that:

$$\Delta_{\min} < z_{jt+1}^{\frac{\tau}{1-\tau}} - \delta z_{jt}^{\frac{\tau}{1-\tau}} < \Delta_{\max}.$$

[35] Given that the index only measures relative price levels, it cannot be used to compare the price level in the same location over time.

Combining the laws of motion of $o_{jt}$ and $y_{jt}$, one obtains the following second order difference equation in $y_{jt}$ :

$$\begin{aligned} y_{jt+1} &= y_{jt}\delta\left(1+F\left(\widehat{v}\right)\left(1-\varphi\right)\right)-\left(1-\varphi\right)\delta^{2}y_{jt-1}F\left(\widehat{v}\right)+ \\ & \left(\frac{\tau}{\overline{w}}\right)^{\frac{1}{1-\tau}}\left(z_{jt+1}^{\frac{\tau}{1-\tau}}-\delta z_{jt}^{\frac{\tau}{1-\tau}}\right)-\delta F\left(\widehat{v}\right)\left(\frac{\tau}{\overline{w}}\right)^{\frac{1}{1-\tau}}\left(z_{jt}^{\frac{\tau}{1-\tau}}-\delta z_{jt-1}^{\frac{\tau}{1-\tau}}\right). \end{aligned}$$

Obviously, since $\varphi<1$ and $F\left(\widehat{v}\right)<1$:

$$y_{jt+1}\leq y_{jt}\delta\left(1+F\left(\widehat{v}\right)\left(1-\varphi\right)\right)+\left(\frac{\tau}{\overline{w}}\right)^{\frac{1}{1-\tau}}\left(\Delta_{\max}-\delta F\left(\widehat{v}\right)\Delta_{\min}\right).$$

Now, assume that

$$\delta\left(1+F\left(\widehat{v}\right)\left(1-\varphi\right)\right)<1. \tag{24}$$

In this case, there exists a $\overline{y}$ such that

$$\overline{y}\leq\left(\frac{\tau}{\overline{w}}\right)^{\frac{1}{1-\tau}}\frac{\left(\Delta_{\max}-\delta F\left(\widehat{v}\right)\Delta_{\min}\right)}{1-\delta\left(1+F\left(\widehat{v}\right)\left(1-\varphi\right)\right)}.$$

Thus, for (23) to hold, it is sufficient that:

$$\Delta_{\min}\geq\left(1-\delta\right)\left[1-\left(1-\varphi\delta\right)F\left(\widehat{v}\right)\right]\frac{\left(\Delta_{\max}-\delta F\left(\widehat{v}\right)\Delta_{\min}\right)}{1-\delta\left(1+F\left(\widehat{v}\right)\left(1-\varphi\right)\right)}.$$

This can be rewritten as:

$$\Delta_{\min}\geq\gamma\Delta_{\max},$$

where

$$\gamma=\frac{\left(1-\delta\right)\left[1-\left(1-\varphi\delta\right)F\left(\widehat{v}\right)\right]}{1-\delta\left(1+F\left(\widehat{v}\right)\left(1-\varphi\right)\right)+\left(1-\delta\right)\left[1-\left(1-\varphi\delta\right)F\left(\widehat{v}\right)\right]\delta F\left(\widehat{v}\right)}.$$

Notice that $\gamma<1$ if $\delta$ is sufficiently low, in which case also condition (24) is satisfied.

Figure 1 − Inflow and Outflow Rates

Correlation Coefficient (Raw Data): 0.66

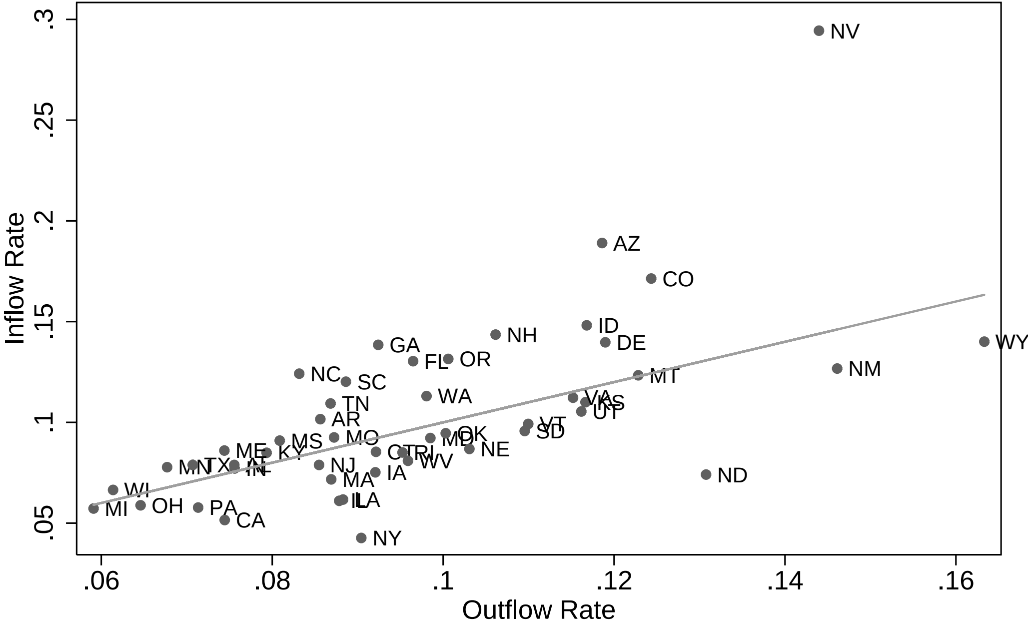
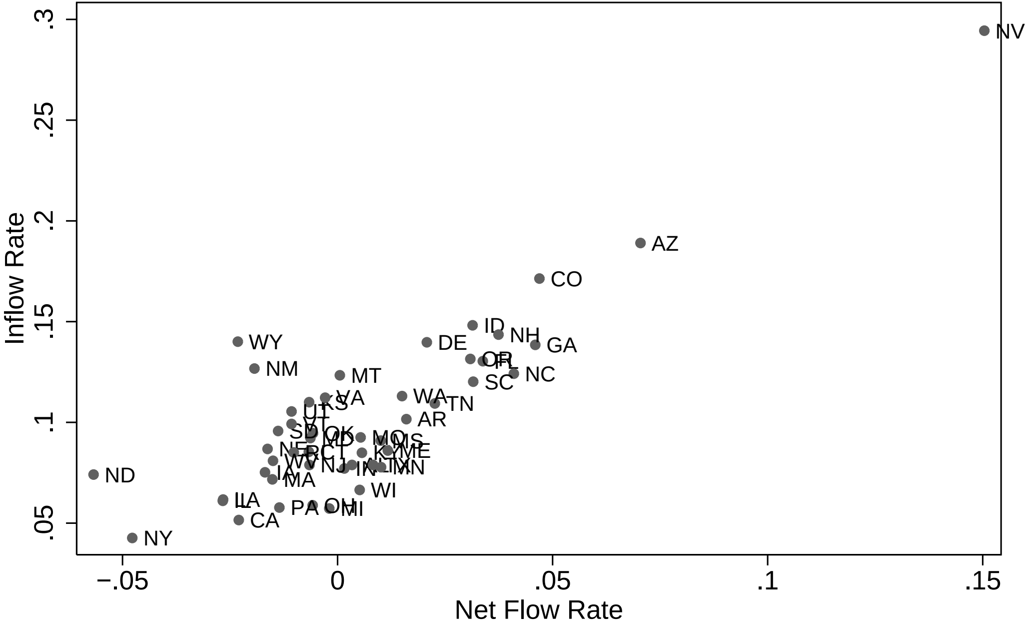


Figure 2 − Inflow and Net Flow Rates

Correlation Coefficient (Raw Data): 0.85



# Figure 3 − Net Flow and Outflow Rates

Correlation Coefficient (Raw Data): 0.16

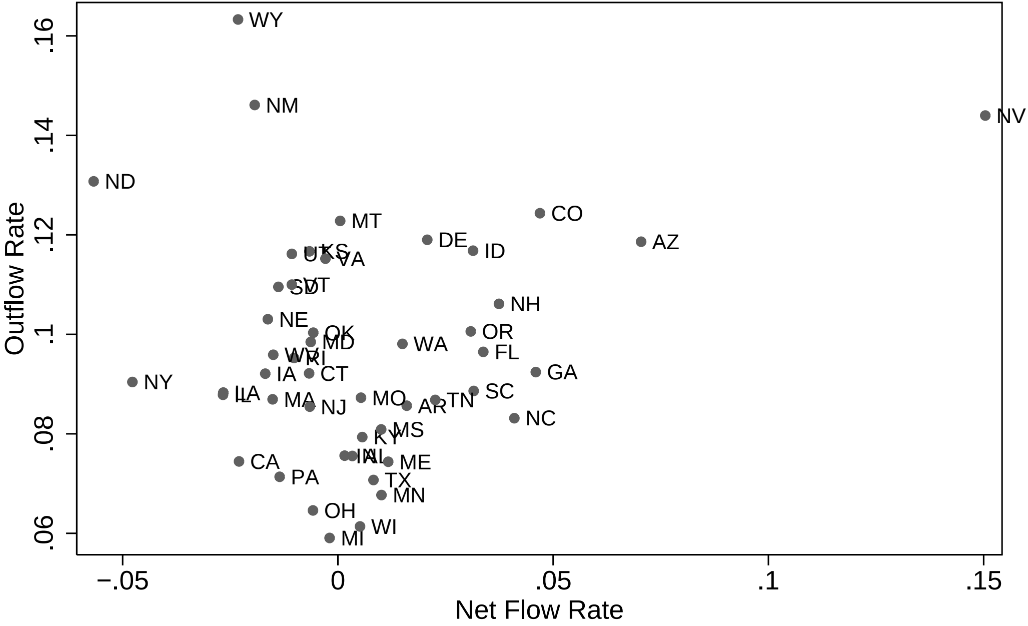


Figure 4 − Net Flow Rates and Average Earnings

Correlation Coefficient: 0.43

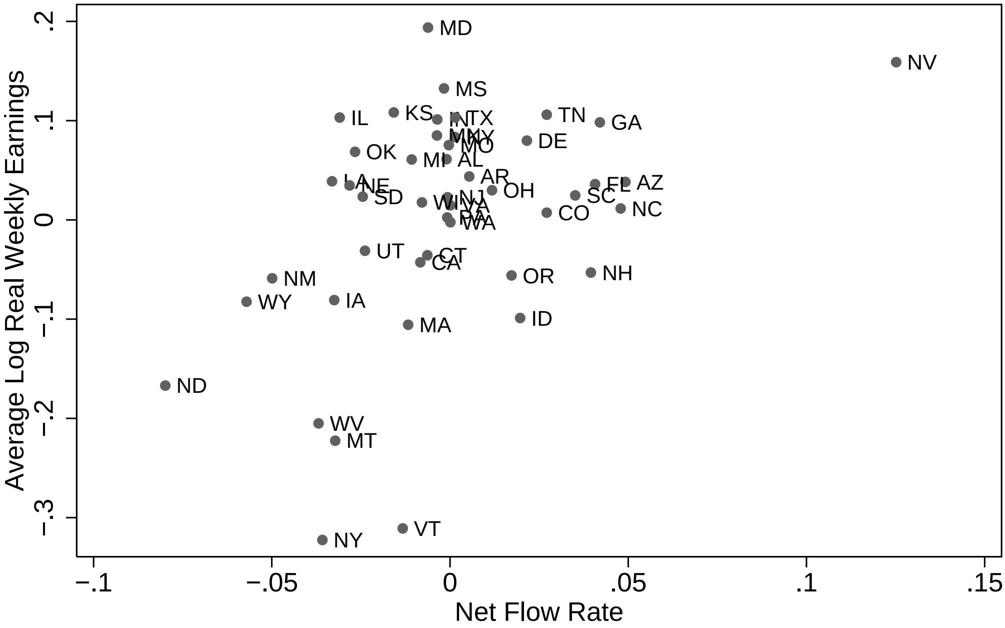
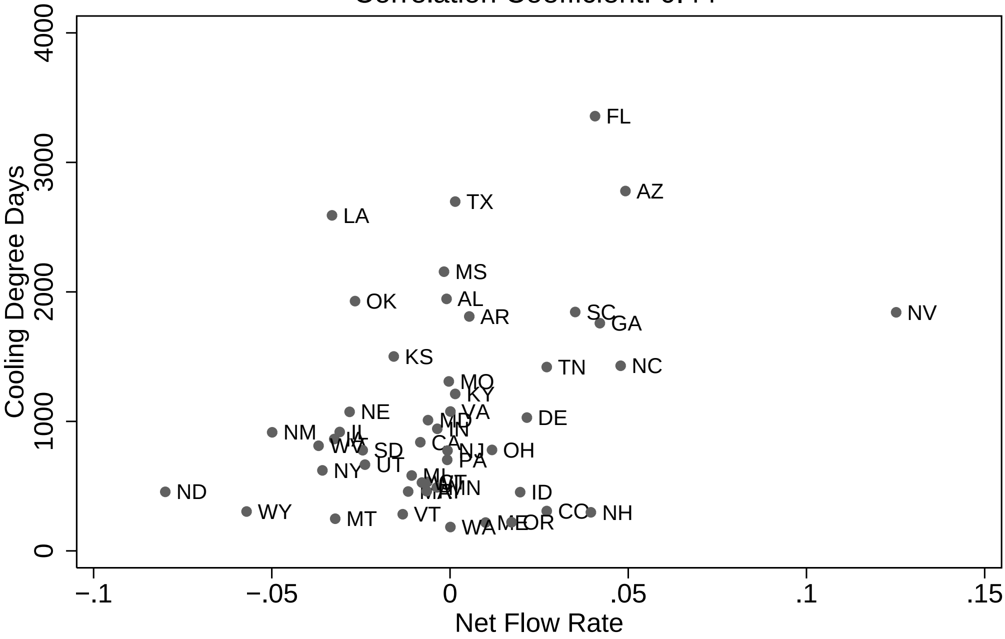
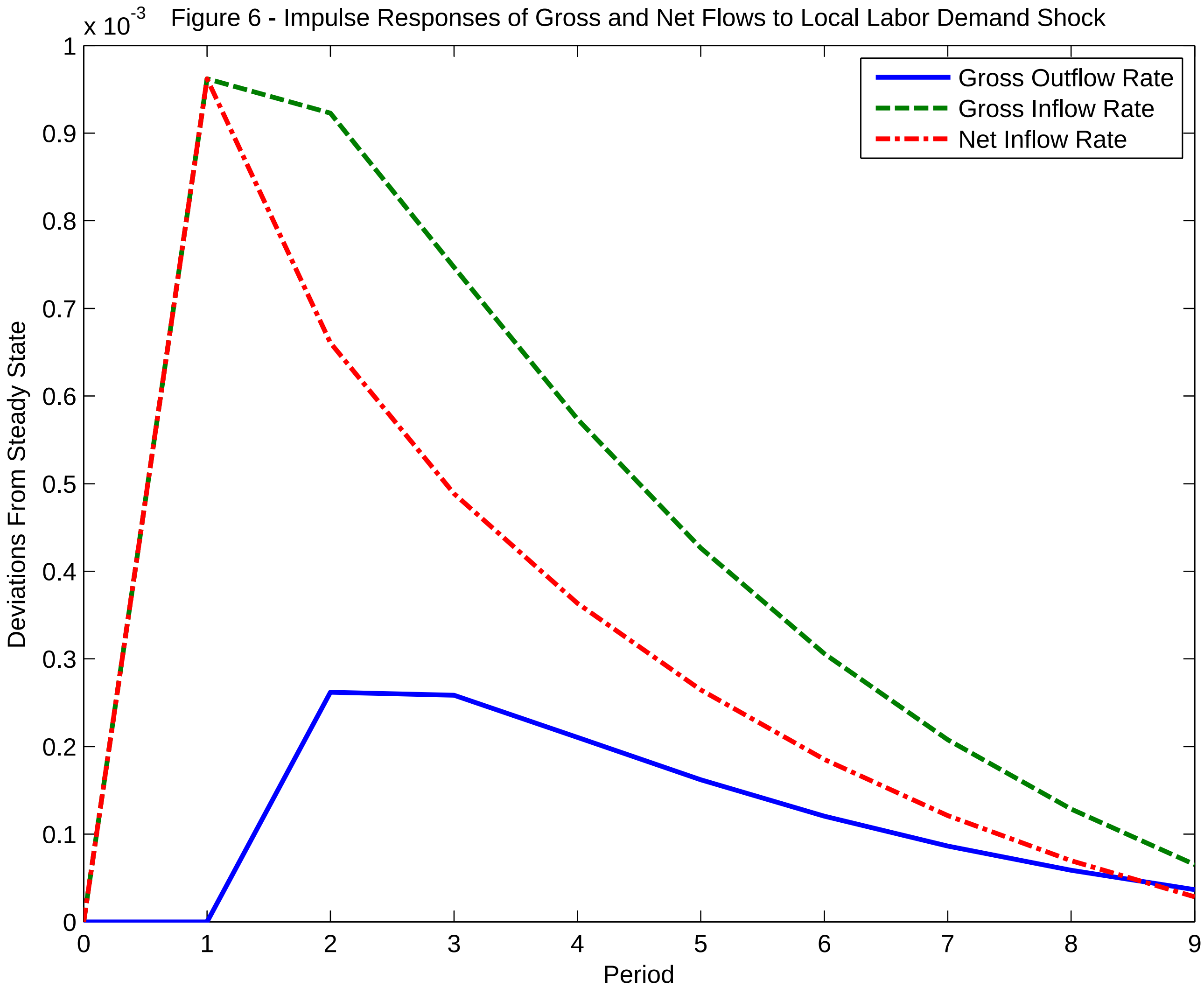


Figure 5 − Net Flow Rates and Hot Weather

Correlation Coefficient: 0.44




x 10^-3
Figure 6 - Impulse Responses of Gross and Net Flows to Local Labor Demand Shock
Gross Outflow Rate
Gross Inflow Rate
Net Inflow Rate
Deviations From Steady State
Period
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
1
2
3
4
5
6
7
8
9

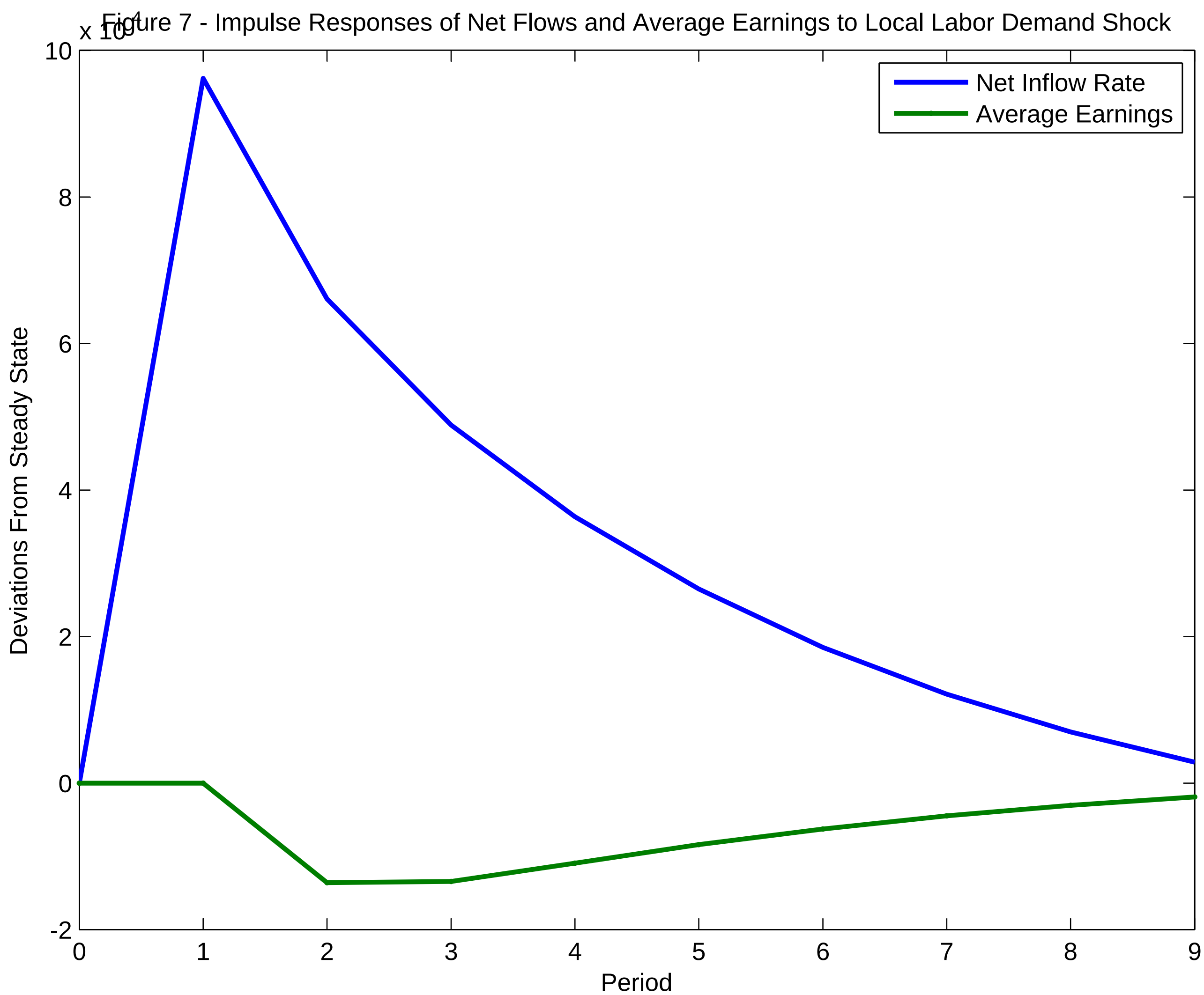


Figure 7 - Impulse Responses of Net Flows and Average Earnings to Local Labor Demand Shock
x 10^-4
Deviations From Steady State
Period
Net Inflow Rate
Average Earnings
10
8
6
4
2
0
-2
0
1
2
3
4
5
6
7
8
9