The Role of Housing in Labor Reallocation*

Morris A. Davis  Jonas D.M. Fisher

University of Wisconsin  Federal Reserve Bank of Chicago
mdavis@bus.wisc.edu  jfisher@frbchi.org

Marcelo Veracierto

Federal Reserve Bank of Chicago
mveracierto@frbchi.org

July 17, 2012

Preliminary and Incomplete

Abstract

Cities experience significant productivity shocks, but population is very slow to respond to these shocks. In light of Blanchard and Katz (1992)’s study showing that migration plays a major role equilibrating regional labor markets, could the slow response of population to productivity shocks help explain why employment remains so low in the US? Our goal in this paper is to develop a framework that can be used to address such a question and use it to assess the importance of local housing markets in determining US population flows. Housing is a natural candidate for slowing population adjustments: houses are difficult to move; they need to be built in order to accommodate population increases; and, relatively large amounts of previously built local housing makes a location more attractive thereby deterring out-migration and encouraging in-migration. We find that these basic characteristics of housing have a large impact on population flows even in the absence of any financial market imperfections.

JEL Classification Numbers: E0, O4, R0

Keywords: Housing, house prices, labor reallocation

*We thank John Kennan, Erzo Luttmer, Robert Shimer and seminar participants at various institutions and conferences for helpful comments. The views expressed herein are those of the authors and do not necessarily represent those of the Federal Reserve Bank of Chicago or the Federal Reserve System.
1 Introduction

Figure 1 shows our point estimates of the response of productivity and population to a standard deviation innovation in productivity for a typical US city over several years. Cities experience significant productivity shocks that are near random walks, but population is very slow to respond to these shocks. In light of Blanchard and Katz (1992)’s study showing that migration plays a major role in equilibrating regional labor markets, could the slow response of population to productivity shocks help explain why employment remains so low in the US? Our goal in this paper is to develop a framework that can be used to address such a question. We use our framework to assess the importance of local housing markets in determining population flows. Housing is a natural candidate for slowing population adjustments: houses are difficult to move; they need to be built in order to accommodate population increases; and, a relatively large stock of previously built housing adds to the attractiveness of a location thereby deterring out-migration and encouraging in-migration. We find that these basic characteristics of housing have a large impact on population flows even in the absence of imperfections in financing housing purchases.

Figure 1: Estimated Response of a City’s Population to a Productivity Shock

In order to provide a guide for selecting the model economy and assessing its empirical plausibility we compile an annual panel data set for 365 Metropolitan Statistical Areas
(MSAs) over the 1985-2009 period. We find that these data have the following key properties: wages and employment fluctuate significantly; employment fluctuates nearly twice as much as population; gross population flows are many times larger times larger than net population flows; house prices are far more volatile than is population; residential construction is extremely volatile; and gross arrivals relative to population are increasing in net migration rates. Substantial fluctuations in wages and employment suggest that cities are subject to significant idiosyncratic shocks. Large gross population flows suggest that moving costs are relatively small. Small net population flows suggests that cities must have a quasi-fixed factor pinning down their sizes. While large fluctuations in house prices support the view that housing is this quasi-fixed factor, the large volatility in their construction seems to contradict this view. The quasi-fixity of housing and large fluctuations in house prices may be due to housing services depending on land, which by nature is in fixed supply. Finally, the increasing arrival rate of migrants as a function of net migration suggests that population flows involve at least some directed search.

Our model is a version of the Lucas and Prescott (1974) islands economy in which islands are interpreted as cities. Local housing services are produced using structures, which can be accumulated at a given location but cannot be moved from that location, and local land that is in fixed supply. Firms produce city-specific goods that are intermediate to final goods production using equipment, land, labor and structures as inputs with a technology that is subject to city-specific idiosyncratic productivity shocks, the unique source of cross-sectional variation. Equipment is freely movable across cities, but labor is costly to move. At the beginning of every period workers must decide whether to stay in their current city or move. Moving to a different city requires the payment of a fixed cost, but workers can move to a city of their choice within the period. Once they have settled in a city a worker must decide whether to become employed or not and the amount of local housing services to consume. To isolate the role of housing’s technological characteristics we assume perfect risk sharing across all agents in the economy.

The model is calibrated to a set of observations that are different from the ones used to assess its empirical plausibility. Some of the calibration targets are closely related to the neoclassical growth model, including the interest rate, labor’s share of income, and several capital:output ratios. The other targets are specific to our environment: the share of land in house prices, the employment:population ratio, construction’s share of total employment, the variance of wages relative to employment, average moving costs, and the gross in-migration rate. A novel feature of our quantitative strategy is that we estimate the idiosyncratic productivity shock process using our panel data on employment and wages thereby pinning down the degree of idiosyncratic variability in the model. The final element of our calibration
is to isolate the degree of substitutability of city-specific intermediate goods. We chose this by matching the model to the empirical cross-section distributions of productivity and population growth, showing that our model is consistent with Zipf’s law for cities.

With only productivity shocks driving cross-sectional variability we find that the model is broadly consistent with the volatility, persistence, and contemporaneous co-movement of the growth of population, employment, wages, house prices and rates of in- and out-migration. Since we expect other idiosyncratic shocks influence city dynamics, for instance shocks to local taxes and spending, this is perhaps surprising. A less stringent test of the model is to compare the dynamic responses of model variables to idiosyncratic productivity shocks with their empirical counterparts we identify in our empirical work. We show that our model does well replicating these dynamics qualitatively, in particular the slow response of population illustrated in Figure 1. In later work we will examine the impact on population flows of eliminating moving costs and allowing structures to be mobile.

Our paper is related to a growing literature which studies geographical dynamics in the context of Lucas-Prescott island economies. Nieuwerburgh and Weil (2010) analyze a model of cities and local housing markets to study the effect of an increase in wage dispersion across cities on the corresponding dispersion of house prices. In contrast to us they abstract from labor supply decisions and moving costs. Lloyd-Ellis and Head (2012) study the interaction among geographical mobility, unemployment and home-ownership in a model with search frictions. They are interested in how the speed with which homeowners can sell their houses influences aggregate unemployment finding this channel to be relatively unimportant in determining the aggregate level of unemployment. They match their model to aggregate means and do not consider the cross-sectional features of the data we emphasize. Coen-Pirani (2010) examines gross and net worker flows, wages and land rents across US states. His focus is the positive correlation between gross inflows and outflows of states. There is no role for housing in his model. While all three of these papers rely on idiosyncratic productivity shocks to drive cross-sectional variation, none of them introduce evidence on the nature of these shocks as we do.

Two other related papers include Alvarez and Shimer (2011) and Kennan and Walker (2011). Similar to us Alvarez and Shimer study a Lucas-Prescott model in which agents can enjoy leisure in the islands where they are located. However, they interpret the islands as industries and use their model to analyze the behavior of wages and unemployment. Their model has no role for housing. Kennan and Walker analyze migration decisions in the face of wage shocks and moving costs. They also abstract from local housing markets, assume undirected search rather than the directed search we consider, and do not study any equilibrium interactions. We use their estimates to calibrate the moving costs in our model.
The rest of the paper is organized as follows. Section 2 describes the data that we use in our analysis. After this we describe the model economy and then how we calibrate the model to U.S. data. In Section 5 we characterize the dynamic response of model variables to productivity shocks. Section 6 discusses how well our model accounts for the within-city dynamics of US cities. Section 7 goes through some experiments with the model to assess the role of housing in labor reallocation. The last section concludes.

2 Empirical Evidence

In this section we introduce the data we use to guide the construction and empirical assessment of our model. We work with an annual panel data set covering 1985 to 2009 of population, net and gross population flows, employment, wages, construction, and house prices for 365 Metropolitan Statistical Areas (MSAs) including 83% of the aggregate population.\footnote{An MSA is a geographical region with a relatively high population density at its core and close economic ties throughout the area determined by measures of commuting patterns. Such regions are not legally incorporated as a city or town would be, nor are they legal administrative divisions like counties or sovereign entities like states. A typical metropolitan area is centered around a single large city that wields substantial influence over the region, \textit{e.g.} Chicago. However, some metropolitan areas contain more than one large city with no single municipality holding a substantially dominant position, \textit{e.g.} the DallasFort Worth metroplex or MinneapolisSaint Paul. With these caveats, for convenience we refer to our MSAs as cities.}

Our focus in this paper is on within-city dynamics. Therefore before analyzing our data we remove both aggregate “time” effects and city-specific “fixed” effects. We remove fixed effects in one of two ways. For population, employment, residential investment, wages per worker and house prices we difference the log of the variable. This transformation removes a level fixed effect. We remove time effects by subtracting from the log growth rate of each city’s variable in a given year the cross-section average log growth rate in that year.

We also study gross and net migration rates and we handle these variables differently. Let $a_{it}$ and $d_{it}$ denote the number of people flowing into and out of city $i$ in year $t$ and $p_{it}$ the population of that city at the end of the same year. For an individual city the gross arrival rate is $a_{it}/p_{it-1}$ and the gross departure rate is $d_{it}/p_{it-1}$. The difference between the gross arrival and departure rates is the net migration rate.

Gross migration rates fluctuate over the business cycle and have been falling over our

\footnote{The appendix describes our data in detail.}
To abstract from these dynamics we subtract from each city gross rate in a year the cross section average of that gross rate in that year. The net migration rates calculated from the resulting gross rates are identical to subtracting from each city’s net migration rate in a year the cross-section average net migration rates in that year. This is how we remove time effects from the migration rates.

Figure 2 contains plots of gross and net migration rates by population decile with time effects. Net migration is essentially independent of city size. This finding reflects Gibrat’s law for cities, that population growth is independent of city size. However, the arrival and departure rates are clearly diminishing in city size. While we think this is an interesting result, ignoring it would confound within-city and across-city variation. We could subtract from each city’s gross migration rates the corresponding time-series average, but this would simultaneously remove a fixed effect from the implied net migration measure which we strongly suspect is not there. To avoid this problem, instead we subtract from each city’s gross migration rates the time-series average of the city’s mean arrival plus departure rate. The corresponding net migration rate for each city is identical to the one calculated after removing time effects alone.

---

\(^2\)See Molloy et al. (2011) and Kaplan and Schultofer-Wohl (2012) for studies of the trend in gross migration rates.
Figure 3 displays the transformed gross migration rates by net migration decile for the full sample of data on migration, 1985-2009. We plot the mean gross or net migration within each decile. Before plotting the transformed gross migration rates we add back the time and city averaged mean of the gross arrival plus departure rates. First, as has been noted by many others, gross migration rates are far in excess of the amount necessary to account for net migration. About 11% of the population either moves in or out of a city in any given year yet net-migration never exceeds 2 and a quarter percent in absolute value. For the sixth decile there is essentially no net migration yet total migrants account almost 12% of the population in these cities. Second, gross arrival rates are monotonically increasing in net migration and gross departure rates are monotonically declining. The rising gross arrival rate suggests that migration involves at least in part directed search. The fact that the two gross migration rates vary similarly suggests that both margins are important when a city’s population adjusts to shocks.

Table 1 shows standard deviations and contemporaneous correlations of several city-level variables. The arrival and departure rates are hat-transformed. Table 1 substantiates most of the claims made about our data in the introduction: wage and employment growth fluctuate a lot; employment growth fluctuates nearly twice as much as population growth; house price growth is far more volatile than is population growth; residential investment growth is extremely volatile. It is also worth noting the large positive correlation of arrival rates and
negative correlations of departure rates with population and employment growth. As was also
evident from Figure 3, both margins of migration appear to adjust to accommodate increases
in the demand for workers. The gross migration rates are weakly negatively correlated. This
finding stands in contrast to the focus of Coen-Pirani (2010) on a positive correlation at the
state level. We believe this difference arises from our removal of city-specific level effects from
our measures of gross migration rates.\(^3\) Finally, notice the large positive correlations of house
price growth with employment and residential investment growth. The latter correlation
partly reflects the way we measure residential investment as house prices multiplied by house
permits.

Table 2: Serial Correlation Within Cities

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.81</td>
<td>0.52</td>
<td>0.82</td>
<td>0.77</td>
<td>0.12</td>
<td>0.73</td>
<td>0.15</td>
</tr>
<tr>
<td>2</td>
<td>0.74</td>
<td>0.29</td>
<td>0.70</td>
<td>0.74</td>
<td>0.10</td>
<td>0.31</td>
<td>0.04</td>
</tr>
<tr>
<td>3</td>
<td>0.67</td>
<td>0.21</td>
<td>0.58</td>
<td>0.67</td>
<td>-0.02</td>
<td>-0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>4</td>
<td>0.63</td>
<td>0.15</td>
<td>0.47</td>
<td>0.60</td>
<td>-0.11</td>
<td>-0.25</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Note: The variables are Population, Employment, Arrival Rate, Departure Rate, Residential Investment, House Prices and Wages. The variables have been transformed as described in the text before calculating the statistics.
investment and wages are the exceptions. House price growth is notably less persistent than population growth or either of the two gross migration rates.

3 The Model Economy

This section describes our model economy. We begin by describing the economy’s agents and their decision problems. After this we describe how the equilibrium can be found by studying a city-planner’s problem wide side conditions. Finally, we describe how to map allocations into wages and house prices. In this version of the model population adjusts on the arrivals margin only. In later drafts we will consider the case where both arrivals and departures adjust.

3.1 The Environment

The economy is populated by a continuum of ex-ante identical agents with names on the unit interval. These agents are initially distributed across a continuum of cities and must decide whether to stay or move. Individual preferences are given by

\[ E \sum_{t=0}^{\infty} \beta^t [\ln c_t - \varphi_t \chi_{n,t} - \psi_t \chi_{m,t} + A \ln \eta_t], \]

where \( c_t, \eta_t, \chi_{n,t}, \chi_{m,t}, \varphi_t \) and \( \psi_t \) denote consumption, housing services, the probability of working, the probability of moving, the disutility of work and the moving cost. The disutility of work \( \varphi_t \) is i.i.d. over time and across individuals according to the distribution function \( \mu_n \) and is realized after moving decisions have been made. The moving cost \( \psi_t \) is i.i.d. according to the distribution function \( \mu_m \) and is realized at the beginning of the period.

Within each city there are four sectors each comprising competitive, price-taking firms endowed with a sector-specific technology. These firms have access to four factors of production: labor, \( n \), which is can be moved across cities at a cost, equipment, \( k \), that is freely mobile, structures, \( h \), which are reproducible but not moveable, and land, \( l \), which is both un-reproducible and un-movable. The final goods sector produces a tradeable commodity using city-specific intermediate goods as inputs. Aggregate supply of the final good is given by

\[ Y_t = \left[ \int_0^1 y_t(j)^\chi \, dj \right]^{\frac{1}{\chi}}, \]

where \( y_t(j) \) is the amount of intermediate good produced by city \( j \) and \( \chi < 1 \). City-\( j \) specific
intermediate goods are produced according to

\[ y_t (j) = s_t n_{yt}^\theta k_{yt}^\gamma h_{yt}^\xi l_{yt}^{1-\theta-\gamma-\xi}, \]

where \( s_t \) is a productivity shock idiosyncratic to the city with transition matrix \( Q \). We have dropped the dependence of the productive factors on the city’s location for simplicity. The parameters \( \theta, \gamma \) and \( \xi \) are all positive and sum to unity so the production function obeys constant returns in the four factors. The first subscript on the factors indicate the sector in which the factors are allocated. The construction sector in city \( j \) produces new structures \( i^h \) according to

\[ i^h_t (j) = n_{ht}^\alpha k_{ht}^\lambda h_{ht}^\kappa l_{ht}^{1-\alpha-\lambda-\kappa}. \]

Finally there is a residential housing sector that combines structures and land to produce housing services according to

\[ \eta_t (j) = h_{rt}^\varsigma l_{rt}^{1-\varsigma}. \]

Structures accumulate at the city level and satisfy

\[ h_{t+1} (j) = (1 - \delta_h) h_t (j) + i^h_t (j) \]

for each city \( j \) where \( h_t (j) \) is the stock of structures in city \( j \) at the beginning of period \( t \) and investment in city \( j \) structures is irreversible, \( i^h_t (j) \geq 0 \). Equipment accumulates in the aggregate as

\[ K_{t+1} = (1 - \delta_k) K_t + I^k_t \]

where \( K \) and \( I^k \) denote the aggregate stock of equipment and investment in new equipment. Equipment is owned by the agents in all cities.

The disutility of work and moving costs are specified as follows. Total disutility of work in a given city depends on the total number of individuals \( p \) and the total number of those individuals who work \( n \). We make the parametric assumption that the total city-level disutility of work is given by

\[ \int_0^{\bar{\varphi}(n/p)} \varphi \mu_n (d\varphi) = \phi \left( \frac{n}{p} \right)^\pi, \]

where \( \bar{\varphi}(n/p) \) satisfies

\[ \frac{n}{p} = \int_0^{\bar{\varphi}(n/p)} \mu_n (d\varphi) \]

and the parameters satisfy \( \phi > 0 \) and \( \pi \geq 1 \). Moving costs are specified in a similar way. For total number of individuals \( x \) in a city and the number of those individuals who move \( m \)
total moving costs are assumed to obey
\[ \int_{-\infty}^{\tilde{\psi}(m/x)} \psi \mu_m (d\psi) = -\psi_2 \ln \left(1 - \frac{m}{x}\right) - \psi_1 \frac{m}{x} \]

where \( \tilde{\psi} (m/x) \) satisfies that
\[ \frac{m}{x} = \int_{-\infty}^{\tilde{\psi}(m/x)} \mu_m (d\psi) \]

and the parameters \( \psi_1 \) and \( \psi_2 \) are both positive. We elaborate on the interpretation of \( \psi_1 \) and \( \psi_2 \) below.

### 3.2 Steady State Equilibrium

We consider a steady state competitive equilibrium with complete markets. Since this is a convex economy with no distortions the welfare theorems apply. As a consequence the equilibrium allocation can be obtained by solving the problem of a social planner that maximizes the expected utility of a representative agent subject to resource feasibility constraints. However it is more useful to characterize the equilibrium allocation as the solution to a series of city social planner’s problems, one for each city, with some side conditions. This approach to studying equilibrium allocations follows Alvarez and Veracierto (2012).

The city planner enters a period with a stock of structures, a population, current and lagged productivity. Denote the planner’s state as \( z = (h, x, s, s_{-1}) \). Taking as given aggregate non-construction output, \( Y \), aggregate consumption, \( C \), the shadow value of equipment, \( r_k \) and the shadow value of movers exclusive of moving costs, \( \Omega \) the planner solves

\[
V(z) = \max \left\{ \frac{1}{\chi} Y^{1-\chi} \left[ sn_y^\theta k_y^\gamma h_s^\xi l_{s-1'}^{1-\gamma-\xi} \right]^{\chi} + CA \ln \left( \frac{h_{s-1'}}{p} \right) p 
- C\phi (n_y + n_h)^{\pi} p^{1-\pi} - r_k (k_y + k_h) 
+ \Omega m - \Omega a + C \left[ \psi_2 \ln(1 - \frac{m}{x}) + \psi_1 \frac{m}{x} \right] x + \beta EV(z') \right\}
\]
subject to

\[ p = x + a - m \]
\[ n_y + n_h \leq p \]
\[ a \geq 0 \]
\[ m \geq 0 \]
\[ h_y + h_r + h_h \leq h \]
\[ l_y + l_r + l_h \leq \bar{l} \]
\[ x' = p \]
\[ h' = (1 - \delta_h) h + n_h^{\alpha_1} k_h^{\lambda_1} l_h^{1 - \alpha - \lambda - \kappa} , \]

where \( m \) is the number of people that move from the city, \( a \) is the number of people that arrive to the city, and \( \bar{l} \) is the exogenous fixed stock of land in the city.

The objective of this optimization problem is to maximize producer surplus. To see this note that the first term is the value of intermediate good production and the second term is the amount of housing services enjoyed in the city. The next five terms comprise the costs to the planner of obtaining this surplus: the disutility of sending the indicated number of people to work (for exogenously fixed shift lengths), the shadow cost of equipment used in the city, and the cost of net migration inclusive of moving workers out of the city. The final term is the discounted continuation value given the updated state vector. The constraints are straightforward too. In order of appearance, they indicate the available population as the sum of the initial population and net in-migration, the available workforce, the constraints that both arrivals and departures must be non-negative, the local resource constraints for structures and land, and the transition equations for population and the stock of structures. Note that depending on the parameters of the utility cost of working and the moving cost, it is possible that there will be unemployed workers and it will always be worth it for the planner to move workers away for sufficiently low moving costs as the cost of replacing these workers is fixed to the planner.

The equilibrium allocations are found by solving the planner’s problem such that the following side conditions are satisfied:

\[ \mu (\mathcal{X}, \mathcal{H}, s', s_{-1}) = \int \{ z: p(z) \in \mathcal{X} \text{ and } (1-\delta_h)h + i^k(z) \in \mathcal{H} \} Q(s'; s, s_{-1}) d\mu \]
\[ Y = \left\{ \int [sn_y^\theta k_y^\gamma h_y^\xi l_y^\varphi] \chi d\mu \right\}^{\frac{1}{\chi}} \]
\[ \int [k_y + k_h] \, d\mu = K \]
\[ C + \delta K = Y \]
\[ \int pd\mu = 1 \]
\[ r_k = \frac{1}{\beta} - 1 + \delta \]

These side conditions are that the steady state distribution of the state vector is reproduced by its endogenous transition equation, the supply of non-construction goods must equal the production of all the cities, the employed equipment must equal its supply, consumption and investment must equal the supply of non-construction final goods, the population allocated to all the cities equals the assumed number of agents, and the rental rate on capital satisfies the intertemporal Euler equation for efficient equipment accumulation. The side conditions are used to solve for the steady state distribution of the state vector and the steady state values for \( Y, K, C, \Omega, \) and \( r_k. \)

### 3.3 Wages and House Prices

Equilibrium prices are straightforward to calculate from the quantity allocations derived from solving the city-planner’s problem with side conditions. Wages in a city are derived from the optimization problem of a firm in the intermediate good sector (equivalently the construction sector) and therefore correspond to the marginal product of labor in the intermediate good sector (construction sector). The rental prices are obtained by considering the optimization problem of the housing service providers:

\[ r_h = CA \frac{S}{h_r} p \]
\[ r_l = CA \frac{(1 - \varsigma)}{l_r} p. \]

Corresponding to these rental prices are the prices of units of structures, \( q^h, \) and land, \( q^l, \) which are derived from the optimization problems of firms in the construction sector and of landlords:

\[ q^h (z) = r_h (z) + (1 - \delta_h) \beta E q^h (z') \]
\[ q^l (z) = r_l (z) + \beta E q^l (z'). \]
We define house prices as the total value of structures and land used to produce housing services per unit of housing services provided. This corresponds to a price of housing per square foot, \( q^{sf} \) under the assumption that every square foot of built housing yields the same quantity of housing services:

\[
q^{sf}(z) = \frac{q^h(z) h_r(z) + q^l(z) l_r(z)}{h_{rt}(z)^\varsigma l_{rt}(z)^{1-\varsigma}}.
\]

4 Calibration

In this section we calibrate the steady state competitive equilibrium described above to U.S. data using a model time period equal to one year. Our calibration has two important characteristics. First, we obtain estimates of the city-specific productivity process via an indirect inference procedure. This means that the amount of volatility in our model is pinned down by our data on productivity. Second, the calibration targets for the remaining parameters involve features of the data that are not primary to our study. So, for instance, we do not choose parameters to fit our estimated response of population to a productivity shock. The response of population to a productivity shock in our model will be the consequence of our estimated productivity process and parameters chosen to fit other features of the data.

We begin by describing how we map our model into the National Income and Product Accounts (NIPA). The model’s stock of non-structure capital \( K \) is identified with the stock of equipment and software. As a consequence, we identify \( \delta_k K \) with private fixed investment in equipment and software. Private investment in structures is measured as \( \int \beta Eq^h(z') i^h d\mu \).

We map model consumption \( C \) to total personal consumption expenditures less imputed housing services. Model output of non-construction final goods, \( Y \), is the sum of consumption, \( C \) and equipment and software investment \( I^k \). We then define the empirical counterpart of \( GDP \) as:

\[
\text{GDP} = Y + \int \left[ \beta Eq^h(z') \right] i^h d\mu,
\]

that is model \( GDP \) is equated with total expenditures on consumption net of housing plus equipment and software investment plus private investment in structures. Observe that we do not include government expenditures and net exports in our definition of \( GDP \) since our model abstracts from them.

In addition to specifying the stochastic process for a city’s productivity we need to find values for 15 parameters:

\( \theta, \gamma, \xi, \alpha, \lambda, \kappa, \varsigma, \delta_k, \delta_h, \beta, A, \pi, \psi_1, \psi_2, \chi \)
These include the factor shares in production, depreciation rates, the discount factor, the coefficient on housing services in agents’ preferences, the parameter determining the elasticity of labor supply at the city level, the parameters underlying the moving costs and the parameter that determines the elasticity of substitution between city-specific intermediate goods.

Values for all but the last four parameters are obtained by reproducing in the model’s steady state (as closely as possible) the same number of aggregate first moments. Seven of these first moments are capital or investment to GDP ratios, including investment in equipment and software (target is 0.11), equipment and software in the non-construction sector (0.60), equipment and software in the construction sector (0.02), private investment in structures (0.11), structures in the non-construction sector (2.54), structures in the construction sector (0.01), structures in the housing sector (1.58). Other moments include the land component value of residential housing from Davis and Heathcote (2007) (0.37), the real interest rate (0.04), labor’s share of income (0.64), the employment to population ratio (0.63) and construction’s share of total employment (0.07).

To identify the magnitude of \( \pi \), which governs the labor supply elasticity at the city level, we exploit the city planner’s first order condition for labor supply. This condition can be written

\[
w = C \phi \pi \left( \frac{n_y + n_h}{p} \right)^{\pi-1}
\]

An implication of this equation is the ratio of standard deviations for wage growth and growth in the employment to population ratio is equal to \( \pi - 1 \). We select \( \pi \) to match the ratio for this variable we estimate with our data. The value we obtain implies a Frisch elasticity of labor supply equal to 1.05.

The moving cost parameters, \( \psi_1 \) and \( \psi_2 \), are identified using evidence on average moving costs and the average gross expansion rate. Specifically, we match the average moving cost per mover relative to average wages to the estimate of this magnitude obtained by Kennan and Walker (2011). Specifically, we require

\[
-C \int \left[ \psi_2 \ln(1 - \frac{m}{x}) + \psi_1 \frac{m}{x} \right] \frac{x}{m} d\mu \int w d\mu = -1.9
\]

to hold. We set the average gross expansion rate equal to its average over cities and dates in

\[\text{4}\]

In most cases we are able to match these moments almost exactly. The exceptions are structures in the non-construction sector and in the residential sector, which in the model have ratios to GDP of 2.24 and 1.5 instead of their targets 2.54 and 1.58.
our sample. Specifically we require

$$\frac{\int m \, d\mu}{\int p \, d\mu} = 0.043.$$ 

The resulting magnitudes of $\psi_1$ and $\psi_2$ imply a cost function that is zero at the origin, falls below zero and then shoots above zero to infinity at the upper extreme of the cost distribution, $\mu_m$.

The final element of calibration is to assign a value to the parameter that determines the elasticity-of-substitution of city-specific intermediate goods, $\chi$, and to specify the stochastic process driving productivity fluctuations. Our calibration of $\chi$ is closely tied to how we go about specifying the productivity process and so we begin by describing this.

Our strategy for estimating the productivity process builds on the idea of indirect inference introduced by Anthony A. Smith (1993). We specify a city-level production function. Based on our model this production function is miss-specified. We identify city-specific productivity by exploiting the implied first-order conditions of an intermediate goods firm subject to the miss-specified technology. From this measure we estimate a stochastic process for productivity growth. The final step is to match this process to the one estimated in the same way from data simulated from our model.

The miss-specified production function is

$$y_t = s_t n_t^\theta k_t^{1-\theta}. $$

Assuming it is perfectly mobile the rental rate of capital, $k_{yt}$, is common to all cities. It follows from the first order conditions associated with the factor input choices of a firm endowed with this technology in any two cities $i$ and $j$ that

$$\frac{w_{it}}{w_{jt}} = \left( \frac{s_{it}}{s_{jt}} \right)^{\frac{\chi}{(1-\theta)\chi}} \left( \frac{n_{it}}{n_{jt}} \right)^{-\frac{1-\chi}{1-(1-\theta)\chi}}. $$

It follows that

$$\Delta \hat{s}_{it} = \frac{1 - (1 - \theta) \chi}{\chi} \Delta \hat{w}_{it} + \frac{\chi}{1 - \chi} \Delta \hat{n}_{it},$$

where $\Delta$ is the first difference operator and the “hats” indicates the variables have been transformed in the same way as their empirical counterparts in Section 2. We identify $\Delta \hat{s}_{it}$ using our data on wages to measure $\Delta \hat{w}_{it}$ and our data on employment to measure $\Delta \hat{n}_{it}$, along with predetermined values for $\chi$ and $\theta$.

When we apply this methodology for plausible values of $\theta$ and $\chi$ we obtain positive serial
autocorrelation coefficients for $\Delta \hat{s}_{it}$. For example for $\chi = 0.9$ and our calibrated labor share $\theta = 0.64$ the serial correlation coefficient, $\rho$, equals 0.19. Serial correlation is zero at longer lags, suggesting that an autocorrelated process in the growth rate of productivity is a good fit for the data. This suggests considering the following process for city-specific productivity:

$$\ln s_{t+1} - \ln s_t = g + \rho (\ln s_t - \ln s_{t-1}) + \varepsilon_{t+1},$$

where $\varepsilon_{t+1} \sim N(0, \sigma^2_\varepsilon)$.

Unfortunately with this specification our estimates of serial correlation for productivity growth imply $\rho > 1$, which makes the process non-stationary. We address this by adopting a reflective barrier for city-specific productivity. Specifically:

$$\ln s_{t+1} = \max \{g + (1 + \rho) \ln s_t - \rho \ln s_{t-1} + \varepsilon_{t+1}, \ln s_{\min}\}.$$

According to this process $s_t$ is reflected at the barrier $s_{\min}$.

The case $\rho = 0$ has been studied thoroughly in the context of cities by Gabaix (1999). In this case $g < 0$ generates a stationary process where the invariant distribution has an exponential tail given by

$$\Pr [s_t > b] = \frac{a}{b^\zeta}$$

for scalars $a$ and $b$. A striking characteristic of cities is that for many countries one finds that $\zeta \simeq 1$. This is called Zipf’s law. For convenience we refer to $\zeta$ as the Zipf’s law coefficient.

The case $\rho > 0$, which applies under our estimates for serial correlation in city-specific productivity growth, has not been studied. Our simulations suggest this process behaves similarly to the $\rho = 0$ case in that the invariant distribution also has an exponential tail. It turns out that a version of Zipf’s law actually holds for our measures of $\hat{s}_{it}$ and so using the reflecting barrier process seems appropriate. We find that the Zipf coefficient for city-specific productivity equals 5.3 at our calibrated value for $\chi$.

We use the drift parameter $g$ and the substitution parameter $\chi$ to match the Zipf coefficients for productivity, $\hat{s}_{it}$, and population, $p_{it}$. For $g = -0.0016$ and $\chi = 0.9$ our model implies Zipf coefficients for productivity and population very similar to the data. With these parameters our model implies Zipf coefficients for population equal to 1.3 and productivity equal to 4.7. We demonstrate our model’s success in generating the cross-section of city populations in Figure 4. The dynamic response of productivity in Figure 1 reflects value of $\rho$ implied by our data with $\chi = 0.9$ and the estimated standard deviation of the productivity innovation $\sigma_\varepsilon = 1.27\%$. 

16
5 Characterization of Model Dynamics

In this section we characterization the model’s predictions for within-city dynamics. We do this by examining a series of impulse response plots calculated by regressing simulated model variables on current and lagged innovations to productivity in each city.

Figure 5 demonstrates that our model successfully reproduces, at least qualitatively, the slow response of population to a productivity shock. The slow response of population is mirrored in the response of the total stock of structures. Not only does the new population need to be accommodated but the larger scale of production justified by the higher level of productivity also requires new structures. This process takes time and limits the expansion of the city.

Figure 6 demonstrates that the growth in structures is obtained by an increase in overall employment but also a shift in the employment share of the construction sector. The shift in resources to the construction sector is mirrored in the response of non-structures capital and land use. A downfall of the model as it now stands is that within the construction sector structures are allocated first to the production sector and away from housing. See Figure 7. With population rising and the stock of structures in housing, housing per person plummets. This is a drawback of the model that we believe is due to assuming a relatively high elasticity of demand for housing in our preference specification. One way to see this is to imagine that agent’s had inelastic demand for housing services. In this case the stock of structures devoted to housing would have to rise if population is to rise. We plan on addressing this issue in
Figure 5: Response of a City’s Population and Structures to a Productivity Shock

Figure 6: Response of a City’s Employment to a Productivity Shock

Figure 8 displays the responses of wages and house prices. Most interesting about this plot
is the (slight) serially correlated growth in house prices. Such dynamics are often viewed as challenging for rational expectations models, e.g. Glaeser (). Here they come about because of the slow response of structures relative to population.
Figure 9 displays the responses of the arrival and departure rates. The model implies that essentially all population adjustment is through arrivals. Departures are constant. This is a direct result of our cost specification; unless at corners it is always optimal to chose the fraction of leavers in a city to minimize costs and there is a unique solution to this minimization problem. We intend to address this downfall of our model in future drafts.

6 Model Evaluation

We now assess the model’s empirical predictions. A traditional way to do this is to compare unconditional model-generated statistics with those we have estimated, such as those in Tables 1 and 2. While this is a worthwhile endeavor (which we will pursue in a later draft), it is really a very strong test of the model. We are confident cities are subject to shocks other than those to productivity and these shocks will impact the variation of our variables of interest. Therefore it is unreasonable to expect our model to reproduce all the unconditional dynamics. Indeed we might consider such a finding to be a drawback of the model.

Instead of studying unconditional dynamics, we focus on conditional dynamics. In particular the dynamic response of variables to productivity shocks. To do this we estimate with our data the dynamic responses of population, employment, house prices, wages, and gross
migration rates (in a later draft we will also examine residential investment.) We estimate these responses as follows. First, from the estimated productivity process we back out a time series of pseudo-innovations to productivity. These are pseudo-innovations because they are based on the miss-specified production function. Second, we estimate a dynamic panel regression with the variable of interest (transformed as described in Section 2) on the left-hand side and current and four lags of the pseudo-innovations on the right-hand side. This delivers the information required to construct impulse response functions for the variables of interest. We compare these to model implied responses by using the identical estimation strategy with simulated data. Note that these are not the same impulse responses described in the previous section which are based on actual productivity innovations. In practice they are very similar, which suggests our identification strategy is successfully capturing the true empirical responses (conditional on our model being correct, of course.)

Figure 10 displays model and estimated responses to a one standard deviation (pseudo-) innovation to productivity along with plus and minus 2 standard error bands for the estimates. We make three observations on this figure. First, our estimation strategy uncovers statistically significant responses (the wage standard error are very small since in this case we are essentially regressing wages on wages.) Second, the model does an excellent job replicating the qualitative pattern of responses. For example, our estimated response of house prices exhibits serially correlated growth as in our model. Third, based on the standard errors our model is still some way from being a quantitative success. Of most concern to us are the counter-factually strong responses of population and employment in the model.

We suspect the source of this quantitative failure is the absence of a departure margin to adjust population and employment in the model. Figure 11 shows what we already knew that almost all population adjustment comes through arrivals in our model and that this response is much too strong. In our model, once the moving cost is paid it is costless for agents’ to allocate to the best cities. In those cities arrivals spike on impact and stay high to accommodate the relatively large (relative to the data) population increases. We suspect that eliminating the moving cost would have essentially no impact on our findings – our’s is a model of net migration, not gross migration. In a later draft we will address this shortcoming by modifying our gross migration technology to accommodate adjustments on both the arrival and departure margins. We expect such a modification to dampen the response of population and employment to a shock.

Finally, observe from Figure 11 that we estimate statistically significant responses of the departure and arrival rates of opposite sign. We find this encouraging for our modeling strategy because it strongly resembles the kind of unconditional dynamics apparent in the gross migration rates displayed in Figure 9.
Figure 10: Model and Estimated Responses to a Productivity Shock

Figure 11: Estimated Response of a City’s Population to a Productivity Shock
7 Measuring the Impact of Housing on Labor Reallocation

This will appear in later drafts when we are more satisfied with the empirical performance of our model.

8 Conclusion

To be completed.

References


A Data Sources

To provide an empirical underpinning to our analysis we construct an annual panel data set of population, employment, population inflows and outflows, residential investment, house prices, and wages of a set of Metropolitan Statistical Areas (MSAs). The mappings of counties to MSAs we use are consistent with the definitions given by the U.S. Office of Management and Budget (OMB) as of December, 2009.\(^5\) OMB currently defines 366 MSAs in the United States. Over our sample period, these MSAs account for about 83 percent of the aggregate population.

For population, employment, wages, and gross migration (inflow- and outflow- rates), we study a balanced panel of 365 MSAs over the period 1985 (the first year of our gross migration data) to 2009. Due to the limited availability of house price data for many MSAs prior to 1990, for house prices and residential investment we study a balanced panel of 172 MSAs over the same 1985-2009 period. We exclude all data from the New Orleans MSA from our study because of the disruption caused by Hurricane Katrina in 2005.

Data on population, employment, and wages are taken from table CA04 of the “Regional Economic Accounts: Local Area Personal Income and Employment” as produced by the Bureau of Economic Analysis (BEA). These data are available at [http://www.bea.gov/iTable/iTable.cfm?ReqID=70&step=1&isuri=1&acrdn=5](http://www.bea.gov/iTable/iTable.cfm?ReqID=70&step=1&isuri=1&acrdn=5).

- The population data (line 20) are mid-year estimates from the Census Bureau.

- For employment, we use “Wage and Salary Employment,” line 7020. These are counts of full- and part- time jobs of salaried employees. The BEA also publishes an estimate of “Total Employment,” which is the sum of wage and salary employment and proprietors employment. The level and change of the two employment series are highly correlated.\(^6\)

- We construct nominal average wage per job, “wages”, as the sum of total wage and salary disbursements, line 50, and supplements to wages and salaries, line 60, divided by wage and salary employment.

We convert the nominal estimate of average wage per job to a real measure by deflating using the CPI “All items less shelter.” The specific CPI series we use is CUSR0000SA0L2, available for download at [http://www.bls.gov/cpi/](http://www.bls.gov/cpi/). We use this series to convert nominal variables to real variables.

For house prices, we use the repeat-sales house price indexes produced by the Federal Home Finance Agency (FHFA).\(^7\) These data are quarterly; for each MSA, we construct an annual estimate as the average of all non-missing quarterly observations. All 366 MSAs have price-index data from 2001 through 2009, but prior to 2001, the time span of coverage varies by MSA. Summarizing: Only 14 MSAs have house price indexes starting in 1976; 130 MSAs


\(^6\)We use wage and salary employment to be consistent with our measure of wages, described next; we do not want to include proprietor’s income as part of wages because some of proprietor’s income represents payments to capital.

starting in 1980; 202 MSAs starting in 1985; 326 MSAs starting in 1990, and so forth. We use the CPI excluding shelter to convert these nominal indexes to real. For 11 MSAs, the FHFA does not report an MSA-level index but rather a set of indexes corresponding to divisions within the MSA. For these MSAs, we set the level of the index for the MSA equal to the average of the reported indexes for the underlying divisions. Because we use an index of house prices, the growth rates of house prices can be compared across MSAs, but levels are not directly comparable without a further adjustment.

To construct an estimate of residential investment, we use data on housing permits and house prices. County-level data on housing permits are available from 1980-2009 as part of the “State of the Cities Data System” (SOCDS) available on the Department of Housing and Urban Development (HUD) Web Site. We aggregate the county-level permits data to the MSA-level. We generate MSA-level indexes of nominal residential investment by multiplying the MSA-level house price indexes described earlier and building permits. We then deflate this estimate using the CPI excluding shelter to produce an index of real residential investment at the MSA-level. As with house prices, the growth rates of the indexes of real residential investment can be compared across MSAs, but the levels cannot.

Finally, we construct data on gross MSA-level population inflows and outflows using county-county migration data based on tax records that is constructed by the Internal Revenue Service (IRS). These data are available annually from 1983 through 1992 at the Inter-University Consortium for Political and Social Research (ICPSR) web site; are available for purchase from the IRS for the 1992 - 2004 period; and are available for free on the IRS web site from 2004 through 2009. The data are annual and cover the “filing year” period, not calendar year. For example, the data for 2007 approximately refer to migration over the period April, 2007 to April, 2008.

For each of the years, the IRS reports the migration data using two files, one for county outflows and one for county inflows, for each county in the United States. Both the inflow and the outflow files report migrants in units of “returns” and in units of “personal exemptions.” According to information from the IRS web site, the returns data approximates the number of households and the personal exemptions data approximates the population. In the data work below, we study the exemptions data.

We define gross inflows into an MSA as the sum of all migrants into any county in that MSA, as long as the inflows did not originate from a county within the MSA. Analogously, we define gross outflows from an MSA as the sum of all migrants leaving any county in that MSA, as long as the migrants did not ultimately move to another county in the MSA.

---

8 The MSAs are Boston, Chicago, Dallas, Detroit, Los Angeles, Miami, New York City, Philadelphia, San Francisco, Seattle, and Washington, DC. Taking Boston as an example, the FHFA reports price indexes for each of the the 4 Metropolitan Divisions of the Boston MSA rather than the Boston MSA itself.

9 The data are available at http://socds.huduser.org/permits/.

10 The value of new housing includes value from residential investment and value from land. Our calculation assumes that land’s share of the value of new housing is constant over time in each MSA.

11 Our procedure for filtering the data described in section 2 is robust to changes in aggregate construction price indexes that do not match changes in the aggregate CPI excluding shelter. The more important assumption we make is that each city has the same growth rate of construction costs. RS Means produce data on construction costs at the MSA level but only for a limited number of MSAs.

12 We exclude the 1983 and 1984 data from our analysis.

exclude people migrating into- and out of the United States. But otherwise, for gross inflows the originating counties are not restricted to be part of one of the 366 MSAs, and for gross outflows the counties receiving the migrants are not restricted to be included in one of the 366 MSAs. Over our sample period, counties inside MSAs slightly increased in population, on-net, relative to counties outside of MSAs.

Define \( a_{it} \) as the number of new entrants to MSA \( i \) during year \( t \) and define \( l_{it} \) as the number of people leaving MSA \( i \) during year \( t \). We compute gross inflow and outflow rates using the IRS data as follows. Define \( \bar{p}_{it} \) as all the people that did not move into or out of MSA \( i \) in year \( t \). We define the beginning of year \( p_{it}^b \) and end of year \( p_{it}^e \) population as

\[
\begin{align*}
p_{it}^b &= \bar{p}_{it} + l_{it} \\
p_{it}^e &= \bar{p}_{it} + a_{it}
\end{align*}
\]

Net migration is therefore \( p_{it}^e - p_{it}^b = a_{it} - l_{it} \). Note that we construct these measures because the IRS data show gains over time to the aggregate population of taxpayers (through foreign in-migration and births in excess of deaths). In other words, the beginning of year population in period \( t + 1 \) is not, by construction, equal to the end of year population in period \( t \).

We define an average population throughout the year as

\[
p_{it} = 0.5 \left( p_{it}^b + p_{it}^e \right)
\]

and compute gross inflow and outflow rates as \( a_{it}/p_{it-1} \) and \( l_{it}/p_{it-1} \). We compute net migration rates as the difference of the inflow and outflow rates.

### B Using the IRS Data

For each of the years 1992-2009, the IRS reports the migration data using 102 Excel files.\(^{14}\) This represents 2 files – one for county outflows and one for county inflows – for each U.S. state and the District of Columbia.

#### B.1 A Description of the Raw County-Level Data

In each of the state inflows files, for every county in that state, the IRS reports the number of people that did not migrate into that county, i.e. lived in that county for two consecutive years. It also reports, for people that migrated into that county, each of the counties sending the migrants and the number of migrants. Not every county sending migrants is reported: All counties sending a relatively small number of migrants are lumped together into an “other county” category. Analogously, in each of the state outflows files, for every county in that state, the IRS reports the population that lived in that county for two consecutive years (i.e., the non-migrants), and, for the people that migrated out of that county, each of the counties receiving the migrants and their number. Like the inflows file, counties receiving only a small number of migrants are lumped together into an “other county” category.

To illustrate the key features of the data, Figures 12 and 13 show the reported experiences

---

\(^{14}\)For the 1985-1991 data, we work directly with ASCII files. The procedures we document are conceptually similar, although many of the specific details are different.
for Callahan county, Texas, in 2007. Callahan is one of the three counties of the Abilene, TX Metropolitan Statistical Area (MSA) - the other two counties are Jones and Taylor. Figure 12 shows the inflow data for Callahan county. When combined, columns A and B show the FIPS (unique identifying) code for Callahan, 48059. Columns C and D combined identify the FIPS code of the sending county with columns E and F providing a description of the sending county and state. Columns G, H, and I show the number of returns and exemptions as well as the adjusted gross income of the migrants. Focusing on column H, row 1, 436 shows the total number (U.S. and foreign) of in-migrants into Callahan (1,099); row 1, 441 shows the total number of people that had a Callahan county address for two consecutive years, the non-migrants (9,602); and rows 1, 442 through 1, 446 show the migrants by county for the major counties – in this case any county sending at least 14 people to Callahan county, TX. 463 people moved from Taylor to Callahan and 40 moved from Jones to Callahan such that 503 of the 1,099 people moving into Callahan county were within-MSA migrants. 596 of the in-migrants into Callahan were from some somewhere else in the United States or another country. We compute the “end of year population” (in this example, end of year 2007) for Callahan county as the sum of the number of non-migrants, row 1, 441, and the total number of in-migrants, row 1, 436.

Figure 13 shows the outflow data for Callahan county for 2007. The data are organized similar to the inflow file, except that columns A and B show the originating county and columns C through I describe the data for the receiving counties. Focusing again on column H, the exemptions data, row 1, 260 shows the total number of migrants out of Callahan county (1,000); row 1, 265 shows the total number of non-migrants, 9,602 – the same as in the inflow file; and, rows 1, 266 through 1, 271 show, for any county receiving at least 26 residents from Callahan, the specific counties receiving residents from Callahan. 434 people moved from Callahan to Taylor (row 1, 266) and 38 moved from Callahan to Jones (row 1, 267), such that 472 of the 1,000 people moving out of Callahan county were within-MSA migrants. 528 of the out-migrants from Callahan were moving somewhere else in the United States or abroad. We compute the “beginning of year” population for Callahan county as the sum of the number of non-migrants, row 1, 265, and the total number of out-migrants from Callahan, row 1, 260.

The data shown in Figures 12 and 13 illustrate a key feature of the migration data: Gross movements of the population in and out of locations (in this case, counties) are an order of magnitude larger than net movements. In the case of Callahan county, TX in 2007, about 10 percent \( \frac{1,050}{9,602 + 1,050} \) of the population turned over in 2007. However, measured as the number of in-migrants less the number of out-migrants, the population of Callahan county only increased by 99 people \( (= 1,099 \text{ in-migrants} - 1,000 \text{ out-migrants}) \), a net increase of less than 1 percent.

B.2 Mapping the Raw County-Level Data to MSA-Level Data

A Metropolitan Statistical Area is defined as a grouping of counties. The United States contains 366 MSAs that account for approximately 83 percent of the population over the 1985-2009 period. The mapping from counties to MSAs is available at the Office of Management and Budget. We use the November, 2008 mapping, available at http://www.census.gov/population/www/metroareas/lists/2008/List4.txt, to map counties to MSAs for every year in our sample. Given this mapping, for each MSA we can track within-
MSA migration, which we define as migration to a different county but in the same MSA, and across-MSA migration, which we define as migration to a county outside the current MSA of residence. For the purposes of this data appendix, define total migration as the sum of within- and across- MSA migration. Continuing with the example of the inflow file for the three counties comprising Abilene, TX: Total migrants into Abilene is defined as the sum of total migrants into each of Callahan, Jones, and Taylor counties; within-MSA migration into to Abilene is defined as the sum of the migration of residents from Jones and Taylor to Callahan, Callahan and Taylor to Jones, and Callahan and Jones to Taylor; and across-MSA migration to Abilene is defined as total migrants into Abilene less within-MSA migration. We use analogous definitions and procedures when working with the outflow files.

As mentioned earlier, the IRS does not always report the number of migrants originating from a county (inflow file) or number of migrants arriving to a county (outflow file). Due to the omission of some detail, the number of within-MSA migrants according to the inflow files does not have to equal the number of within-MSA migrants according to the outflow files: This can occur whenever a relatively small county in an MSA is lumped into the “Other” category. To illustrate magnitudes, a list of these discrepancies between 1995 and 2007 is given in Table 3. This table shows that these discrepancies occur very infrequently, and typically the discrepancy in terms of the number of migrants is quite small. Whenever there is a discrepancy between the inflow and outflow files of the number of within-MSA migrants, we use the maximum.

\footnote{For example, row 1, 447 of Figure 12 reports the number of migrants into Callahan county from “Other Flows - Same State.”}
<table>
<thead>
<tr>
<th>Year</th>
<th>CBSA</th>
<th>MSA</th>
<th>Inflow File</th>
<th>Outflow File</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995-1996</td>
<td>11100</td>
<td>Amarillo, TX (MSA)</td>
<td>6806</td>
<td>6776</td>
</tr>
<tr>
<td></td>
<td>16860</td>
<td>Chattanooga, TN-GA (MSA)</td>
<td>7193</td>
<td>7166</td>
</tr>
<tr>
<td></td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>16879</td>
<td>16785</td>
</tr>
<tr>
<td></td>
<td>41660</td>
<td>San Angelo, TX (MSA)</td>
<td>42</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>46660</td>
<td>Valdosta, GA (MSA)</td>
<td>971</td>
<td>1042</td>
</tr>
<tr>
<td>1996-1996</td>
<td>11100</td>
<td>Amarillo, TX (MSA)</td>
<td>7545</td>
<td>7509</td>
</tr>
<tr>
<td></td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>17775</td>
<td>17769</td>
</tr>
<tr>
<td></td>
<td>41660</td>
<td>San Angelo, TX (MSA)</td>
<td>58</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>46660</td>
<td>Valdosta, GA (MSA)</td>
<td>2098</td>
<td>2046</td>
</tr>
<tr>
<td>1997-1998</td>
<td>11100</td>
<td>Amarillo, TX (MSA)</td>
<td>7204</td>
<td>7210</td>
</tr>
<tr>
<td></td>
<td>31140</td>
<td>Louisville-Jefferson County, KY-IN (MSA)</td>
<td>21700</td>
<td>21386</td>
</tr>
<tr>
<td></td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>18002</td>
<td>16823</td>
</tr>
<tr>
<td></td>
<td>41660</td>
<td>San Angelo, TX (MSA)</td>
<td>48</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>45820</td>
<td>Topeka, KS (MSA)</td>
<td>1989</td>
<td>2026</td>
</tr>
<tr>
<td></td>
<td>46660</td>
<td>Valdosta, GA (MSA)</td>
<td>1617</td>
<td>1679</td>
</tr>
<tr>
<td>1998-1999</td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>19232</td>
<td>19238</td>
</tr>
<tr>
<td></td>
<td>41660</td>
<td>San Angelo, TX (MSA)</td>
<td>76</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>46660</td>
<td>Valdosta, GA (MSA)</td>
<td>1546</td>
<td>1523</td>
</tr>
<tr>
<td>1999-2000</td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>19554</td>
<td>19578</td>
</tr>
<tr>
<td></td>
<td>41660</td>
<td>San Angelo, TX (MSA)</td>
<td>42</td>
<td>116</td>
</tr>
<tr>
<td></td>
<td>46660</td>
<td>Valdosta, GA (MSA)</td>
<td>1483</td>
<td>1437</td>
</tr>
<tr>
<td>2000-2001</td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>19457</td>
<td>19382</td>
</tr>
<tr>
<td></td>
<td>41660</td>
<td>San Angelo, TX (MSA)</td>
<td>60</td>
<td>71</td>
</tr>
<tr>
<td>2001-2002</td>
<td>11100</td>
<td>Amarillo, TX (MSA)</td>
<td>7454</td>
<td>7506</td>
</tr>
<tr>
<td></td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>18895</td>
<td>19007</td>
</tr>
<tr>
<td>2002-2003</td>
<td>10900</td>
<td>Allentown-Bethlehem-Easton, PA-NJ (MSA)</td>
<td>10071</td>
<td>9134</td>
</tr>
<tr>
<td></td>
<td>11100</td>
<td>Amarillo, TX (MSA)</td>
<td>7470</td>
<td>7526</td>
</tr>
<tr>
<td></td>
<td>37980</td>
<td>Philadelphia-Camden-Wilmington, PA-NJ-DE-MD (MSA)</td>
<td>112771</td>
<td>112877</td>
</tr>
<tr>
<td></td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>19568</td>
<td>17380</td>
</tr>
<tr>
<td></td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>19712</td>
<td>19681</td>
</tr>
<tr>
<td>2004-2005</td>
<td>11100</td>
<td>Amarillo, TX (MSA)</td>
<td>7980</td>
<td>8031</td>
</tr>
<tr>
<td></td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>20279</td>
<td>20228</td>
</tr>
<tr>
<td></td>
<td>41660</td>
<td>San Angelo, TX (MSA)</td>
<td>34</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>47020</td>
<td>Victoria, TX (MSA)</td>
<td>747</td>
<td>629</td>
</tr>
<tr>
<td>2005-2006</td>
<td>10500</td>
<td>Albany, GA (MSA)</td>
<td>3607</td>
<td>3562</td>
</tr>
<tr>
<td></td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>19917</td>
<td>19842</td>
</tr>
<tr>
<td></td>
<td>41660</td>
<td>San Angelo, TX (MSA)</td>
<td>209</td>
<td>128</td>
</tr>
<tr>
<td>2006-2007</td>
<td>11100</td>
<td>Amarillo, TX (MSA)</td>
<td>7836</td>
<td>7726</td>
</tr>
<tr>
<td></td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>18399</td>
<td>18352</td>
</tr>
<tr>
<td></td>
<td>41660</td>
<td>San Angelo, TX (MSA)</td>
<td>182</td>
<td>83</td>
</tr>
<tr>
<td>2007-2008</td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA (MSA)</td>
<td>18715</td>
<td>18606</td>
</tr>
<tr>
<td></td>
<td>41660</td>
<td>San Angelo, TX (MSA)</td>
<td>63</td>
<td>52</td>
</tr>
</tbody>
</table>
### Figure 12: Screen Shot from 2007-2008 IRS Texas Inflow File

<table>
<thead>
<tr>
<th>FIPS Code</th>
<th>Texas</th>
<th>Migration into</th>
<th>Migration from</th>
<th>Number of returns</th>
<th>Number of exemptions</th>
<th>Aggregate adjusted gross income</th>
</tr>
</thead>
<tbody>
<tr>
<td>1436 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Tot Mig-US &amp; For</td>
<td>523</td>
<td>1,099</td>
<td>19,275</td>
</tr>
<tr>
<td>1437 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Tot Mig-US</td>
<td>513</td>
<td>1,079</td>
<td>19,991</td>
</tr>
<tr>
<td>1438 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Tot Mig-Same St</td>
<td>425</td>
<td>305</td>
<td>15,981</td>
</tr>
<tr>
<td>1439 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Tot Mig-Diff St</td>
<td>88</td>
<td>174</td>
<td>3,019</td>
</tr>
<tr>
<td>1440 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Tot Mig-Diff St</td>
<td>16</td>
<td>26</td>
<td>284</td>
</tr>
<tr>
<td>1441 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>4,246</td>
<td>9,602</td>
<td>181,140</td>
</tr>
<tr>
<td>1442 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>227</td>
<td>463</td>
<td>7,947</td>
</tr>
<tr>
<td>1443 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>17</td>
<td>47</td>
<td>592</td>
</tr>
<tr>
<td>1444 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>16</td>
<td>35</td>
<td>721</td>
</tr>
<tr>
<td>1445 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>15</td>
<td>23</td>
<td>465</td>
</tr>
<tr>
<td>1446 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>10</td>
<td>14</td>
<td>199</td>
</tr>
<tr>
<td>1447 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>143</td>
<td>336</td>
<td>6,217</td>
</tr>
<tr>
<td>1448 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>128</td>
<td>194</td>
<td>3,019</td>
</tr>
<tr>
<td>1449 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>14</td>
<td>16</td>
<td>284</td>
</tr>
</tbody>
</table>

### Figure 13: Screen Shot from 2007-2008 IRS Texas Outflow File

<table>
<thead>
<tr>
<th>FIPS Code</th>
<th>Texas</th>
<th>Migration from</th>
<th>Migration into</th>
<th>Number of returns</th>
<th>Number of exemptions</th>
<th>Aggregate adjusted gross income</th>
</tr>
</thead>
<tbody>
<tr>
<td>1260 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Tot Mig-US &amp; For</td>
<td>503</td>
<td>1,000</td>
<td>16,286</td>
</tr>
<tr>
<td>1261 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Tot Mig-US</td>
<td>503</td>
<td>1,000</td>
<td>16,286</td>
</tr>
<tr>
<td>1262 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Tot Mig-Same St</td>
<td>426</td>
<td>367</td>
<td>13,933</td>
</tr>
<tr>
<td>1263 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Tot Mig-Diff St</td>
<td>76</td>
<td>131</td>
<td>2,350</td>
</tr>
<tr>
<td>1264 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Tot Mig-Diff St</td>
<td>6</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>1265 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>4,246</td>
<td>9,602</td>
<td>181,140</td>
</tr>
<tr>
<td>1266 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>221</td>
<td>434</td>
<td>8,923</td>
</tr>
<tr>
<td>1267 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>18</td>
<td>32</td>
<td>638</td>
</tr>
<tr>
<td>1268 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>17</td>
<td>38</td>
<td>526</td>
</tr>
<tr>
<td>1269 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>13</td>
<td>21</td>
<td>302</td>
</tr>
<tr>
<td>1270 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>12</td>
<td>26</td>
<td>374</td>
</tr>
<tr>
<td>1271 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>11</td>
<td>17</td>
<td>608</td>
</tr>
<tr>
<td>1272 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>136</td>
<td>244</td>
<td>4,583</td>
</tr>
<tr>
<td>1273 48</td>
<td>TX</td>
<td>059 000 000</td>
<td>Callahan County Non-Migrants</td>
<td>74</td>
<td>133</td>
<td>2,350</td>
</tr>
</tbody>
</table>