Gross Migration, Housing and Urban Population Dynamics*

Morris A. Davis
Rutgers University
morris.a.davis@rutgers.edu

Jonas D.M. Fisher
Federal Reserve Bank of Chicago
jfisher@frbchi.org

Marcelo Veracierto
Federal Reserve Bank of Chicago
mveracierto@frbchi.org

March 18, 2016

Abstract

Cities experience near random walk productivity shocks, yet population is slow to adjust to them. We use a calibrated dynamic general equilibrium model of cities to disentangle the influence of two natural sources of this slow adjustment: inelastic local housing supply and the intrinsic costs of migration. Our model incorporates a new theory of migration and reproduces patterns we uncover in a panel of US cities. We find that the intrinsic costs of migration are the prime driver of slow population adjustments while housing plays a more limited role. Furthermore we find that migration costs can explain persistent urban decline.

JEL Classification Numbers: E0, O4, R0
Keywords: Gross migration, housing, urban population dynamics, persistent urban decline, house prices

*An earlier draft of this paper circulated under the title "The Role of Housing in Labor Reallocation." We thank Marco Bassetto, Jeff Campbell, Gilles Duranton, Edward Glaeser, John Kennan, Erzo Luttmer, Fernando Parro, Nicolas Petrosky-Nadeau, Robert Shimer and seminar participants at various institutions and conferences for helpful comments. The views expressed herein are the those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Chicago or the Federal Reserve System.

1. Introduction

It is well-documented that US states, cities and counties all experience highly persistent population growth.[1] In this paper we introduce complementary evidence: at the annual frequency cities experience significant, random-walk-like productivity shocks, yet population is very slow to adjust to them. Without any impediments to moving, population should adjust quickly as workers migrate to the more productive cities.[2] There must be substantial barriers to population re-allocation. This paper seeks to understand and quantify these barriers.

There are two natural barriers to rapid population movements: adjustment costs in migration and housing. Using state-level data, Blanchard and Katz (1992) argue that persistent labor demand shocks are accommodated over the long run through migration. Kennan and Walker (2011) use individual-level data to estimate interstate migration propensities as a function of relative wages and intrinsic migration costs. These propensities imply very slow population adjustments to permanent wage shocks. If migration is important, then the decisions over whether to leave one's current location and if so where to move to should be crucial to understanding population dynamics.

Migration may be important, but the speed with which it occurs does not depend on its intrinsic costs alone. New workers require housing so the speed of population adjustments is limited by the supply of local housing. Housing is naturally slow to adjust to shocks because it is immobile across cities and the competing uses of local labor imply increasing marginal construction costs. Glaeser and Gyourko (2005) identify an additional channel. They argue that housing impedes population declines when a large, durable and immobile stock makes housing relatively cheap and discourages people to leave a city.

[1] For examples, see Blanchard and Katz (1992), Glaeser, Scheinkman, and Shleifer (1995) and Rappaport (2004).

[2] See Turek (1985) for evidence that net migration rather than births and deaths is the dominant source of regional variation in population growth in the US.

This paper quantifies the impact of migration and housing on urban population dynamics by constructing a quantitative dynamic stochastic general equilibrium model of cities and analyzing it within the context of a panel dataset comprising 365 US cities over the period 1985-2007. Blanchard and Katz (1992), Glaeser and Gyourko (2005) and Kennan and Walker (2011) do not isolate the role of housing from the intrinsic costs that slow down migration. Our framework makes this possible.

Measuring migration's intrinsic costs is key to our findings. Our measurement is based on a striking empirical regularity we uncover regarding the relationship between gross and net migration. Specifically, the annual arrival rate of a city (total arrivals divided by the initial population) is essentially *linearly* increasing in its corresponding net migration rate (total arrivals minus total departures, divided by the initial population.) Linearity places strong restrictions on the nature of migration decisions and delivers a path to measuring their costs via the magnitude of the slope coefficient. Given its importance to our analysis, we provide corroborating evidence for linearity. In particular we verify the prediction that if the annual arrivals-versus-net relationship is linear then a similar linear relationship with the *identical* slope coefficient should hold for migration rates measured over longer horizons.

The slope of the arrivals-versus-net relationship lies between 0 and 1 which suggests that a city's population adjusts by movements in arrivals and departures in opposite directions. Such a negative correlation might seem surprising given Kennan and Walker (2011)'s finding that recent arrivals are more likely to leave a city than long-time residents and Coen-Pirani (2010)'s demonstration that this phenomenon can induce a positive correlation. However a positive labor demand shock makes incumbents who are on the margin of leaving a city less likely to do so while at the same time it attracts migrants, which induces a negative correlation.

Consistent with the operation of this mechanism, we estimate that a city's arrival rate rises and departure rate falls after it receives a positive productivity shock. Our empirical findings lead us to abstract from the recent arrival phenomenon in our analysis.

To disentangle the roles of housing and migration in slow population adjustments we integrate a new model of inter-city migration inspired by our empirical findings into an otherwise familiar generalization of the neoclassical growth model with endogenous local housing and labor supply. We include endogenous labor supply because employing more of the existing population is a natural substitute for migration to accommodate higher labor demand. Indeed Blanchard and Katz (1992) identify this as one of the main channels through which a state's employment responds in the years immediately following a labor demand shock.

Our modeling of migration shares some features with the literature, but in contrast it is consistent with the strong restrictions implied by our linearity finding. As in Kennan and Walker (2011) the decision to leave a city is driven by idiosyncratic shocks to workers' tastes for their current city relative to others. Conditional on moving, a worker's location choice includes a directed component because empirically arrival rates are increasing in net migration (otherwise arrival rates would be independent of net migration.) Costly directed migration captures the many ways workers find new cities to live and work, including via informal contacts between friends and family, professional networks, specialized firms like head-hunters, advertising that promotes cities as desirable places to live and work, firms' human resource departments, and via recruiting by workers whose primary responsibility is some other productive activity.[3]

We measure the roles of housing and migration in explaining slow population

[3]Our model abstracts from the pervasive decline in gross migration that has been modeled by Karahan and Rhee (2014) and Kaplan and Schulhofer-Wohl (2015). It is an open question whether the mechanisms we emphasize would interact with those that have been used to explain the trend.

adjustments by studying a calibrated version of our model. The measurement of migration costs suggested by our linearity finding is key to our calibration procedure. In addition we use microeconomic evidence on migration; the population distribution across cities; new evidence on city productivity dynamics; a new estimate of a city's wage-elasticity of labor supply; and aggregate ratios familiar from other studies based on the neoclassical growth model.

To build confidence in our calibrated model we compare it to dynamic responses to productivity shocks of a city's population, gross migration, employment, wages, residential investment and house prices we estimate from our panel data. Despite calibrating our model to different information, the model does well along these dimensions and importantly it reproduces the slow population dynamics. We also find that the model is broadly consistent with key unconditional features of the data such as relative volatility, contemporaneous co-movement, and persistence at the annual frequency and the longer run persistence that has been emphasized in the literature. While being driven only by productivity shocks, our model generates a substantial fraction of the variation we find in our data.

Having established the empirical relevance of our model we use it to quantify the influence of housing and migration on population adjustments. We find that workers' costs of finding new cities are the predominant source of slow population adjustments suggesting that they represent significant barriers to population reallocation. In the absence of migration frictions, adjustment costs in housing do lower the amplitude and increase the persistence of population's response to a productivity shock. However, if migration costs are already present, adding adjustment costs for housing does little to influence population dynamics.

We also use our model to examine the determinants of persistent urban decline. In our data the populations of many cities decline throughout the sample period. Glaeser and Gyourko (2005, pp. 368–369) dismiss a role for migration in explaining this behavior. We find that declining cities also experience declining relative productivity, so our model can in principle account for persistent urban decline. In fact it does, through the slow response of population to past declines in productivity. This suggests that costly migration could be a major factor in explaining persistent urban decline even though, as emphasized by Glaeser and Gyourko (2005), arrival rates remain high in declining cities.

Our model is a dynamic version of the classic Roback (1982) and Rosen (1979) model of cities with housing and costless mobility. Because it is static, the Roback-Rosen model does not address migration or population adjustments to shocks. Van Nieuwerburgh and Weil (2010) introduce dynamics to this framework. Their model has implications for net migration, but not gross flows. Coen-Pirani (2010) also constructs a dynamic Roback-Rosen model. He studies gross worker flows among US states in a search and matching framework borrowed from studies of gross worker flows among firms. We find stark differences between gross worker flows between firms compared to those between cities and so we adopt a different approach to modeling migration. Rappaport (2004) studies a two-city equilibrium model with quadratic adjustment costs in net migration, but he does not identify the size of the costs. While the reduced form of our model similarly involves quadratic adjustment costs in net migration, we are able to identify the magnitude of these costs by exploiting our linearity finding.

We contribute to an extensive empirical and theoretical microeconomic literature on migration, surveyed by Greenwood (1997) and Lucas (1997). Kennan and Walker (2011) is an important recent contribution to this literature. Their analysis is centered around the migration choice problem of an individual and abstracts from housing. In contrast we include housing and equilibrium interactions. Mangum (2015) studies housing and migration in the context of a choice problem like Kennan and Walker (2011)'s. While Mangum (2015) focuses on similar issues, he models the housing and labor markets as reduced forms and does not investigate the relationship between between gross and net migration that is central to this paper.

The housing boom and bust in the 2000s prompted a literature that quantifies how housing frictions impede geographical labor reallocation. For example, Karahan and Rhee (2013), Lloyd-Ellis and Head (2012) and Nenov (2015) study how house price collapses limit labor reallocation through disincentives to migrate arising from home ownership and local search frictions. We abstract from these factors. Nenov (2015) and Karahan and Rhee (2013) model gross migration, but they do not address the linear relationship between gross and net migration.[4]

The paper is organized as follows. We begin by describing the new empirical evidence that guides our analysis. Next we introduce our model of migration within a static setting to highlight its key mechanisms. We then embed this static model within a dynamic stochastic general equilibrium setting; describe how we calibrate this model; establish the empirical relevance of the calibrated model; and quantify the roles of housing and migration implied by the calibrated model. The last section concludes by discussing some drawbacks to our analysis and directions for future research.

2. Empirical Evidence

We analyze an annual panel data-set covering 1985 to 2007 that includes population, net and gross migration, employment, wages, residential construction, and house prices for 365 Metropolitan Statistical Areas (MSAs) comprising about 83% of the aggregate population. An MSA is defined as a region with a rela-

[4] Ferreira, Gyourko, and Tracy (2012), Modestino and Dennett (2013) and Schulhofer-Wohl (2012) investigate the effects of housing-related financial frictions on mobility. We abstract from financial frictions in this paper.

tively high population density and close economic ties throughout as measured by commuting patterns. While they sometimes include multiple legally incorporated cities, for convenience we refer to MSAs as cities. Cities are a natural unit of analysis as they represent geographically distinct labor markets. Appendix A describes the sources of our data and Appendix B discusses our empirical findings in more detail.

2.1. Gross and Net Migration

We calculate annual city-to-city migration rates using data on tax returns from the Internal Revenue Service (IRS). IRS data are unique in covering a large number of US cities. The Current Population Survey and the American Community Survey are widely used to construct annual state-level migration rates, but small samples substantially limit their coverage of cities. We show in Appendix B that our key findings for cities extend to the state-level using all three data-sets. That the state-level results are robust to the data source is noteworthy given the drawbacks of the IRS data highlighted in Kaplan and Schulhofer-Wohl (2015).

Define a city's one year arrival rate as the total number of people that move to a city within a year divided by its beginning-of-year population. The corresponding one year departure rate is defined similarly in terms of people leaving a city. Gross migration rates fluctuate over the business cycle and have been falling over our sample period.[5] To abstract from these dynamics we subtract from a city's annual gross rate the corresponding cross-section average. We define a city's net migration rate as the difference between these two gross rates so that it too is free of a time effect.

Figure 1 displays gross and net migration rates by population decile with the time effects removed for all city-year observations. Net migration is seen to be

[5]See Saks and Wozniak (2011) for evidence on the cyclicality of gross migration and Molloy, Smith, and Wozniak (2011) and Kaplan and Schulhofer-Wohl (2015) for evidence of its trend.

unrelated to city size. This confirms Gibrat's law for cities holds in our dataset. However, gross migration clearly declines with city size. As far as we know this is a new finding. While worthy of further study, its presence conflates cross-city variation with the within-city dynamics that are our focus and so we abstract from it. We do this by subtracting from each year's arrival and departure rate the sample average gross migration rate for that city. This removes city fixed effects in gross migration without affecting net migration rates, which Gibrat's law implies are not present.

Figure 1: Gross and net migration rates by population decile

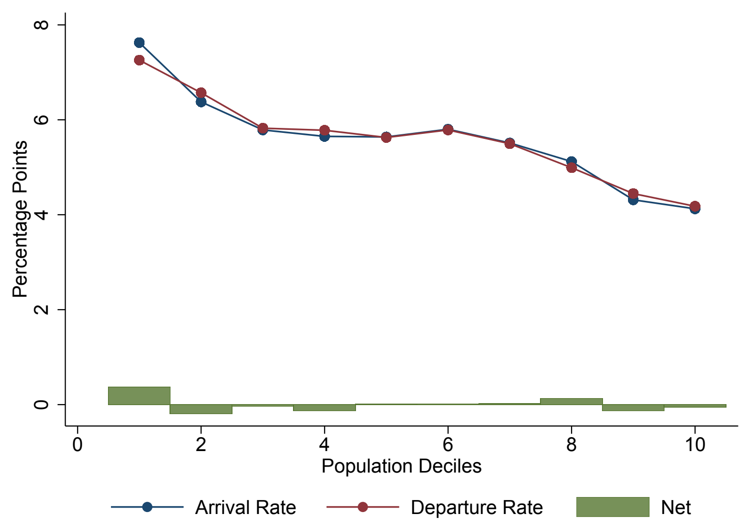
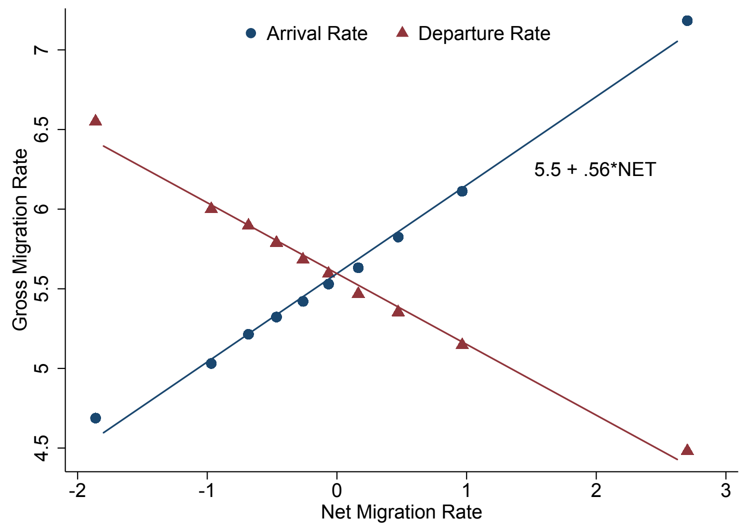


Figure 2 displays mean gross migration rates versus mean net migration rates for each net migration decile, after removing both time and fixed effects and adding the corresponding unconditional mean. We see that the arrival (departure) rate is linearly increasing (decreasing) in net migration with the individual observations lining up almost exactly on the regression lines. The arrival rate being increasing in net migration is *prima facie* evidence of directed migration since otherwise they would be independent. The linearity finding contrasts sharply with the behavior of firm-worker flows found by Davis, Faberman, and

Figure 2: Gross migration rates by net migration



Haltiwanger (2006). Hires are flat near zero for negative net flows and linearly increasing for positive net flows; separations are the mirror image. These differences between worker-city and worker-firm flows suggest models designed to explain the latter are invalid in our context – we require a different theory.

Since linearity is central to our modeling and measurement it is important to have some corroborating evidence for it. We provide this in Appendix B by showing that a prediction of linearity holds in our data. The prediction is that if gross and net migration rates are linearly related at the annual frequency, then they must be linearly related for multi-year migration rates as well and that the multi-year slope coefficients will be *identical* to their one-year counterparts.

Figure 2 reflects that arrival and departure rates are negatively correlated. This finding emerges from addressing city fixed effects in gross migration, since, as Figure 1 shows, the correlation is positive otherwise. Coen-Pirani (2010) reports that state-level gross migration rates are positively correlated. However, we show in Appendix B that after we remove fixed effects state-level gross migration display a very similar pattern to that found in Figure 2.

2.2. Responses of Population and Gross Migration to TFP Shocks

We estimate dynamic responses of city-level variables to local productivity shocks by exploiting the first order conditions of competitive final and intermediate good producers in the model described in Section 4.[6] There are N cities that each produce a distinct intermediate good used as an input into the production of final goods. Intermediate goods in city i at date t are produced according to

$$y_{it} = s_{it} n_{y,it}^{\theta} k_{y,it}^{\gamma}, \tag{1}$$

where s_{it} is exogenous total factor productivity (TFP), $n_{y,it}$ is employment, $k_{y,it}$ is equipment, $\theta, \gamma > 0$, and $\theta + \gamma \leq 1$. The output of the final good at date t, Y_t, is produced using inputs of the city-specific intermediate goods according to

$$Y_t = \left[\sum_{i=1}^{N} y_{it}^{\chi} \right]^{\frac{1}{\chi}}, \tag{2}$$

where $\chi \leq 1$.

Our measurement of city-specific TFP relies on the following definition. For any variable v_{it}:

$$\hat{v}_{it} \equiv \ln v_{it} - \frac{1}{N} \sum_{j=1}^{N} \ln v_{jt}. \tag{3}$$

Subtracting the mean value of $\ln v_{jt}$ in each period eliminates variation due to aggregate shocks, allowing us to focus on within-city dynamics. We assume that equipment is perfectly mobile so that its rental rate is common across cities. It then follows from the first order conditions of final and intermediate good producers that

$$\Delta \hat{s}_{it} = \frac{1 - \gamma\chi}{\chi} \Delta \hat{w}_{it} + \frac{1 - \theta\chi - \gamma\chi}{\chi} \Delta \hat{n}_{y,it}, \tag{4}$$

[6] Lloyd-Ellis, Head, and Sun (2014) estimate responses of city-level variables to personal income shocks identified using a panel VAR.

where Δ is the first difference operator and w_i denotes the wage in city i. Applying the first difference operator addresses non-stationarity over the sample period.[7]

Given values for χ, θ and γ and data on wages and employment we use equation (4) to measure $\Delta\hat{s}_{it}$, the growth rate of city-specific TFP.[8] Below we calibrate θ and γ using traditional methods and find a value for χ to match the model to Zipf's law. With calibrated values $\chi = .9$, $\theta = .66$ and $\gamma = .235$ we estimate an AR(1) in $\Delta\hat{s}_{it}$ with a statistically significant auto-correlation coefficient equal to .24. Wooldridge (2002)'s test of the null hypothesis of no first order serial correlation in the residuals has a p-value of .28, confirming that this specification is a good fit for the data.[9]

Let e_{it} denote the residual from the TFP growth regression. We obtain the response to a TFP shock of variable $\Delta\hat{x}_{it}$ using the coefficients $b_1, b_2, \ldots, b_{10}$ from the following panel regression:

$$\Delta\hat{x}_{it} = \sum_{l=1}^{10} b_l e_{it-l+1} + u_{it} \tag{5}$$

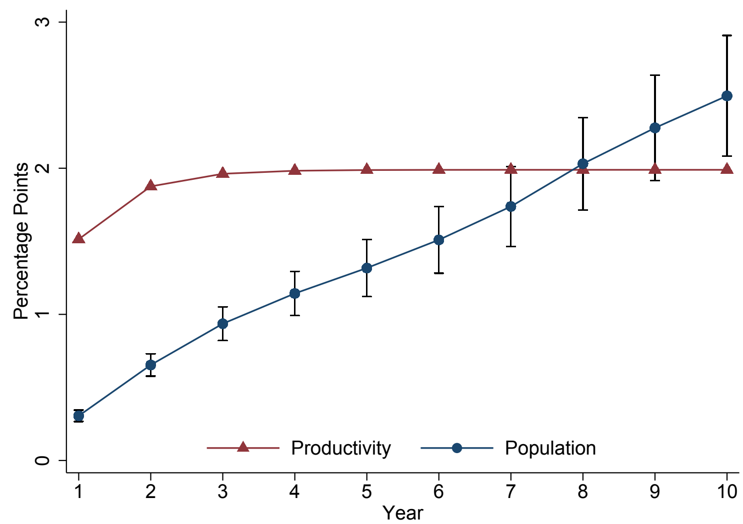
where u_{it} is an orthogonal error term. The response of $\hat{x}_{it}$ is obtained by summing the estimated coefficients appropriately. For the gross migration rates we replace $\Delta\hat{x}_{it}$ with the rates themselves (net of time and fixed effects) in (5) and identify the dynamic responses with the estimated coefficients directly.

[7] It also removes any fixed effects. However, as in Gabaix (1999), our model addresses the cross-section of city populations without appealing to fixed effects.

[8] Total employment in our model includes workers that produce houses and intermediate goods, but we use total employment as our empirical measure of $n_{y,it}$ because we cannot separate out housing employment. Construction employment is a small fraction of total employment so we obtain almost identical results when we exclude total construction employment for the cities it is available.

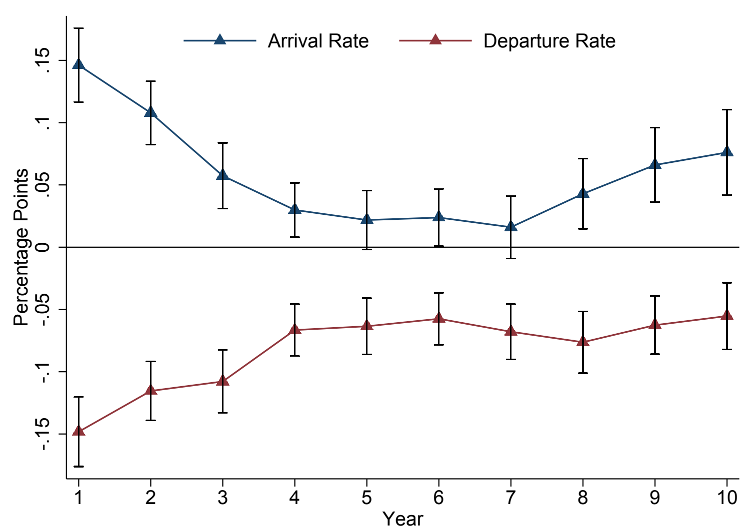
[9] A version of equation (4) can be derived using Davis, Fisher, and Whited (2014)'s agglomeration specification. In that case the coefficients on wage and employment growth include the parameter that governs agglomeration's effect on local productivity. When we re-estimate the TFP process using Davis et al. (2014)'s estimate of this parameter the auto-correlation coefficient and the innovation standard deviation are similar.

Figure 3: Responses of TFP and population



Note: Point estimates along with 2 standard error bands.

Figure 4: Responses of arrival and departure rates



Note: Point estimates along with 2 standard error bands.

Figure 3 and Figure 4 display percentage point deviations of population and gross migration to one standard deviation shocks to TFP growth. Figure 3 establishes the claim made in the introduction that TFP responds much like a random walk, rising quickly to its new long run level, while population responds far more slowly, still growing ten years after the shock. The population response contrasts sharply with Blanchard and Katz (1992) who estimate a reduced form vector auto-regression and find that state populations converge to their new long run level within 7 years after a labor demand shock. Figure 4 corroborates Figure 2 by showing that population adjustments involve responses of arrival and departure rates in opposite directions of similar magnitudes. These responses reinforce the view that improving local prospects could change a city's population both by encouraging locals not to move, which lowers the departure rate, and by attracting new workers, which increases the arrival rate.

3. Modeling Migration

We now introduce our theory of migration by studying a simple, static model which abstracts from housing, equipment, and labor supply. This model is useful for developing intuition about migration choices; describing how we can reproduce the relationships depicted in Figure 2; and establishing the importance of modeling gross migration for understanding urban population dynamics. The main results extend to the model we use for our quantitative analysis.[10]

3.1. A Static Model

The economy consists of a large number of geographically distinct cities with initial population x. In each city there are firms which produce identical, freely tradeable consumption goods with the technology sn^θ, where s is a city-wide TFP shock, n is labor and $0 < \theta < 1$. There is a representative household with a unit continuum of members that are distributed across city types $z = (s, x)$ according to the measure μ. Each household member enjoys consumption, C, and supplies a unit of labor inelastically. After the TFP shocks have been realized, but before

[10] See the technical appendix available at `http://morris.marginalq.com/research.html` for a detailed analysis of the static and quantitative models.

production takes place, the household decides how many of its members leave each city and how many of those chosen to leave move to each city. Once these migration decisions have been made, production and consumption occur.

The departure decision is based on each household member receiving a *location-taste* shock ψ, with measure μ_l, that subtracts from their utility of staying in the city in which they are initially located. This kind of shock appears in the migration models of Kennan and Walker (2011), Karahan and Rhee (2013) and Nenov (2015). To match the empirical relationship between gross and net migration we assume the distribution of individual location-taste shocks in a city of type z satisfies:

$$\int_{-\infty}^{\bar{\psi}(l(z)/x)} \psi d\mu_l = -\psi_1 \frac{l(z)}{x} + \frac{\psi_2}{2} \left(\frac{l(z)}{x} \right)^2$$

where $\psi_1, \psi_2 \geqslant 0$ and $\bar{\psi}\left(l(z)/x\right)$ is defined by $l(z)/x = \int_{-\infty}^{\bar{\psi}(l(z)/x)} d\mu_l$.

This quadratic specification is U-shaped starting at the origin which means that the first workers to leave a city are those for whom leaving raises their utility. As more people leave the remaining inhabitants are those who have a strong preference for staying. These features are consistent with Kennan and Walker (2011)'s evidence that individuals who move receive substantial non-pecuniary benefits and that non-movers would find it extremely costly if they were forced to move. For example, many individuals move to be near family members or find it very costly to move because they are already near family members. Subject to these shocks, the household determines how many of its members from each city must leave to find new cities.

When deciding where to send the leavers the household knows the distribution of city types μ but not the location of any specific type z. However, it can find a particular type through directed migration by obtaining a *guided trip*. We choose a functional form for producing guided trips to match the evidence on gross and net migration. By giving up u units of utility each household member can produce $\sqrt{2}A^{-1/2}u^{1/2}$ guided trips to her initial location where $A \geqslant 0$. Therefore, to attract $a(z)$ workers to a city of the indicated type the household must incur a total utility cost of $(A/2)\left(a(z)/x\right)^2 x$. Costly guided trips encompass the myriad of ways in which workers find a particular city to live and work.

If a household member does not obtain a guided trip it can migrate to another city using *undirected migration*. Specifically, by incurring a utility cost τ a leaver is randomly allocated to another city in proportion to its initial population. Including undirected migration captures the idea that choosing to move to a particular city is often the outcome of idiosyncratic factors other than wages or housing costs that are difficult to model explicitly, such as attractiveness of amenities and proximity to family members.[11] Furthermore, it is natural to let people move to a location without forcing them to find someone to guide them.

We characterize allocations in this economy by solving the following planning problem:

$$\max_{\substack{\{C,\Lambda,a(z),\\ l(z),p(z)\}}} \left\{ \ln C - \int \left[\frac{A}{2}\left(\frac{a\left(z\right)}{x}\right)^2 x + \left(-\psi_1 \frac{l\left(z\right)}{x} + \frac{\psi_2}{2}\left(\frac{l\left(z\right)}{x}\right)^2 \right) x \right] d\mu - \tau\Lambda \right\} \tag{6}$$

subject to

$$p\left(z\right) \leq x + a\left(z\right) + \Lambda x - l\left(z\right), \forall\, z \tag{7}$$

$$\int \left[a\left(z\right) + \Lambda x\right] d\mu \leq \int l\left(z\right) d\mu \tag{8}$$

$$C \leq \int sp\left(z\right)^{\theta} d\mu \tag{9}$$

[11] Kennan and Walker (2011) include both undirected and directed migration. It is undirected because to learn a location's permanent component of wages workers have to migrate there. It is directed because workers retain information about locations to which they have previously migrated and include this information in their current migration decision along with expectations about locations they have not visited already.

and non-negativity constraints on the choice variables.[12] The variable Λ is the fraction of the household that engages in undirected migration. Since these workers are allocated to cities in proportion to their initial populations, Λ also corresponds to the share of a city's initial population that arrives through undirected migration. Constraint (7) states that population in a city is no greater than the initial population plus arrivals through guided trips and undirected migration minus the number of workers who migrate out of the city. Constraint (8) says that total arrivals can be no greater than the total number of workers who migrate out of cities and (9) restricts consumption to be no greater than total production, taking into account that each individual supplies a unit of labor inelastically, $n(z) = p(z)$, $\forall z$.

3.2. The Gross Migration Margins

It is necessary to include frictions on both gross migration margins to match the evidence depicted in Figure 2. Suppose $A = 0$ so that guided trips can be produced at no cost, but that household members continue to be subject to location-taste shocks, $\psi_1, \psi_2 \geqslant 0$. Then it is straightforward to show

$$\frac{a\left(z\right)}{x} = \max\left\{\frac{p\left(z\right) - x}{x} + \frac{\psi_1}{\psi_2}, 0\right\} \text{ and } \frac{l\left(z\right)}{x} = \max\left\{\frac{\psi_1}{\psi_2}, -\left(\frac{p\left(z\right) - x}{x}\right)\right\}.$$

Observe that as long as the net population growth rate, $(p(z) - x)/x$, is not too negative, the planner sets the departure rate, $l(z)/x$ at the point of maximum benefits, ψ_1/ψ_2, and adjusts population using the arrival rate, $a(z)/x$, only. In this situation the departure rate is independent of net population adjustments, contradicting Figure 2. Now suppose that there are no location-taste shocks,

[12] See Appendix C for one possible decentralization.

$\psi_1 = \psi_2 = 0$, but it is costly to create guided trips, $A > 0$. In this case we find

$$\frac{a(z)}{x} = \max\left\{\frac{p(z) - x}{x} - \Lambda, 0\right\} \text{ and } \frac{l(z)}{x} = \max\left\{-\left(\frac{p(z) - x}{x} - \Lambda\right), 0\right\}.$$

The planner always goes to a corner: when net population growth is positive the departure rate is set to zero, and when net population growth is negative the arrival rate is set to zero. Clearly the relationship between gross and net migration in this situation also contradicts Figure 2. Therefore to be consistent with Figure 2 it is necessary to include costs on both gross migration margins.

For the model to be consistent with Figure 2, the number of workers leaving a city and the number arriving to the same city using guided trips must both be strictly positive, $l(z) > 0$ and $a(z) > 0$. The reason we require $l(z) > 0$ is that gross out-migration is always positive in Figure 2. The reason we require $a(z) > 0$ is that otherwise there would be intervals of net migration in which arrival rates are constant, equal to Λ, which is also inconsistent with Figure 2. Therefore, unless otherwise noted, from now on we assume that $a(z) > 0$ and $l(z) > 0$.

Combining the first order conditions for Λ and $a(z)$ we obtain

$$\tau = A \int a(z)\, d\mu.$$

This equation describes the trade-off between using directed and undirected migration. The marginal cost of raising the fraction of household members engaged in undirected migration is equated to the average marginal cost of allocating those household members using guided trips. The averaging reflects the fact that undirected migration allocates workers in proportion to each city's initial population.

The first order conditions for $a(z)$ and $l(z)$ imply that

$$A\frac{a\left(z\right)}{x}=\psi_1-\psi_2\frac{l\left(z\right)}{x}.$$

Intuitively, migration out of a city increases to the point where the marginal benefits of doing so (recall that the location-taste shocks initially imply benefits to leaving a city) are equated with the marginal cost of attracting workers into the city.

As long as $a\left(z\right)>0$ and $l\left(z\right)>0$ population adjustments are independent of the undirected migration decision. Undirected migration is determined by the solution to $\tau=d\Phi(\Lambda)/\Lambda$. In the model we use in our quantitative analysis arrivals are set to zero in especially undesirable cities. Still, for most cases arrivals are strictly positive so that undirected migration is essentially irrelevant for our results. This is a useful property given that there is little evidence on the magnitude of undirected migration.

3.3. Connecting Figure 2 to Population Adjustments

The planner's first order conditions reveal how gross migration relates to net migration. From the first order conditions for $a(z)$ and $l(z)$ and the population constraint, (7), it is straightforward to show that

$$\frac{a\left(z\right)}{x}+\Lambda=\frac{\psi_1}{A+\psi_2}+\frac{A}{A+\psi_2}\Lambda+\frac{\psi_2}{A+\psi_2}\left(\frac{p\left(z\right)-x}{x}\right). \quad (10)$$

The arrival rate is an affine function of the net migration rate $\left(p(z)-x\right)/x$ with the linear coefficient satisfying $0<\psi_2/(\psi_2+A)<1$. The departure rate is also is an affine function of the net migration rate with the linear coefficient satisfying $-1<-A/(\psi_2+A)<0$. These relationships establish that gross migration in the model can be made consistent with Figure 2. This result is the underlying reason for our specifications of the location-taste shocks and the production of guided trips. Clearly, the relationship between gross and net migration depicted in Figure 2 places strong restrictions on the nature of migration costs.

The speed of population adjustments is directly connected to the gross migration choices. This can be seen by substituting for $a(z)$ and $l(z)$ in the original planning problem using their relationships to net migration described above, which simplifies it to

$$\max_{\{p(z),\Lambda\}} \left\{ \ln\left(\int sp(z)^{\theta} \, d\mu \right) - \int \left[\Phi(\Lambda) + \frac{1}{2} \frac{A\psi_2}{A+\psi_2} \left(\frac{p(z)-x}{x} \right)^2 x \right] d\mu - \tau\Lambda \right\}$$

subject to:

$$\int p(z) \, d\mu = \int x d\mu \tag{11}$$

with non-negativity constraints on the choice variables and where $\Phi(\Lambda)$ is a quadratic function in Λ involving the underlying parameters ψ_1, ψ_2 and A. In deriving this simplified planning problem we have used the fact that (7) and (8) reduce to (11) and that this constraint holds with equality at the optimum. Similarly we have used (9) to substitute for consumption in the planner's objective function.

When the planning problem is written in this way we see that population adjustments do not depend directly on $a(z)$ and $l(z)$. Nevertheless modeling these decisions is crucial for understanding population dynamics because the coefficient that determines the speed of population's adjustment to shocks, $A\psi_2/(A+\psi_2)$, involves parameters governing them. Also notice that the reduced form costs of adjusting population are quadratic. This is a direct consequence of specifying the location-taste shocks and guided trip technology to reproduce Figure 2. In other words Figure 2 implies quadratic adjustment costs in net population adjustments.

4. The Quantitative Model

This section describes our quantitative model. We employ a differentiated goods version of the neoclassical growth model that incorporates the gross migration environment described above as well as local residential construction. In principle, both elements could play an important role shaping population dynamics. The role of gross migration was explained in detail in the previous section, but the potential role of housing is more straightforward: since housing is valued by individuals, adjustment costs in the construction sector are likely to be reflected in population dynamics.[13] Because changes in labor supply are a natural alternative to migration to accommodate labor demand shocks the model includes a labor supply decision as well. After describing the model environment we characterize its unique stationary competitive equilibrium as the solution to a representative city planning problem with side conditions.

4.1. The Environment

As before the economy consists of a continuum of geographically distinct locations called cities that are subject to idiosyncratic TFP shocks. Cities are distinguished by their stock of housing, h, initial population, x, and the current and lagged TFP, s and s_{-1}.[14] The measure over the state vector $z = (h, x, s, s_{-1})$ is given by μ.

Within cities there are three production sectors corresponding to intermediate goods, housing services and construction. The representative firm of each sector maximizes profits taking prices as given. Intermediate goods are distinct to a city and imperfectly substitutable in the production of the freely tradeable final goods

[13] This point is demonstrated in Appendix D.

[14] We assume every city has access to the same quantity of developable land and abstract from topographical and regulatory variation across cities. As documented by Saiz (2010) and Gyourko, Saiz, and Summers (2008) in reality there is considerable heterogeneity along these dimensions which can be viewed as affecting the land available for development. However, our analysis focuses on the average response of cities to TFP shocks and not the distribution of these responses across cities.

non-durable consumption and durable equipment. The technologies for producing intermediate and final goods are identical to those underlying our estimates of TFP, described in equations (1) and (2).[15] Housing services are produced by combining residential structures with land, b_r, according to $h^{1-\zeta}b_r^{\zeta}$, $0 < \zeta < 1$. Following the convention that the prime symbol denotes next period's value of a variable, residential structures in a city evolve as

$$h' = (1 - \delta_h)\, h + n_h^{\alpha} k_h^{\vartheta} b_h^{1-\alpha-\vartheta}, \tag{12}$$

where the factor shares are restricted to $\alpha, \vartheta > 0$ and $\alpha + \vartheta < 1$, and $0 < \delta_h < 1$ denotes housing's depreciation rate. The last term in (12) represents housing construction. Local TFP s does not impact residential construction, reflecting our view that residential construction productivity is not a major source of cross-city variation in TFP. Equation (12) embodies our assumptions that residential structures are immobile, durable and costly to build quickly. The latter follows because residential construction requires local labor and land which have alternative uses in intermediate goods production and housing services. We assume that equipment is homogenous and perfectly mobile.

There is an infinitely lived representative household that allocates its unit continuum of members across the cities. One period corresponds to a year. The household faces the same migration choices described in Section 3, but being infinitely lived it takes into account the effects of current migration decisions on its members' allocation across cities in future periods. In particular, it is now bound by the constraint

$$x' = p \tag{13}$$

in each city where p continues to denote the post-migration population of a

[15]Equations (1) and (2) are written in terms of the location of a city, indexed by i, but here it is convenient to index them by the type of the city as represented by its state vector z.

city. The household's members have logarithmic preferences for consumption and housing services in the city in which they are located.[16] Each period, after the migration decisions have been made, but before production and construction take place, individual household members receive a labor disutility shock φ with measure μ_n. Similar to our treatment of migration costs we make a parametric assumption for the average disutility of working. Specifically, if the household decides n of its members in a city will work these costs are specified as

$$\int_{-\infty}^{\bar{\varphi}(n/p)} \varphi d\mu_n = \phi \left(\frac{n}{p} \right)^{\pi},$$

where $\phi > 0$, $\pi \geq 1$ and $\bar{\varphi}(n/p)$ is defined by $n/p = \int_0^{\bar{\varphi}(n/p)} d\mu_n$ The parameter π governs the elasticity of a city's labor supply with respect to the local wage.

4.2. Stationary Competitive Equilibrium

This model has a unique stationary competitive equilibrium. Since it is a convex economy with no distortions, the welfare theorems apply and so the equilibrium allocation can be obtained by solving the problem of a planner that maximizes the expected utility of the representative household subject to the resource feasibility constraints. However, it is more convenient to characterize the equilibrium allocation as the solution to a representative city-planner's problem with side conditions. This approach to studying equilibrium allocation follows Alvarez and Shimer (2011) and Alvarez and Veracierto (2012).

The city planner enters a period with the state vector z. Taking as given aggregate output of final goods, Y, the marginal utility of consumption, λ, the shadow value of adding one individual to the city's population exclusive of the arrival and departure costs, $\lambda\eta$, the shadow value of equipment, λr_k, the arrival

[16]We assume logarithmic preferences for housing to be consistent with the evidence reported by Davis and Ortalo-Mangé (2011) that housing's share in household expenditures is roughly constant across cities.

rate of workers through undirected migration Λ, and the transition function for TFP, $Q\left(s';s,s_{-1}\right)$, the representative city planner solves

$$V\left(z\right)=\max_{\substack{\{n_y,n_h,k_y,k_h,\\ h',b_r,b_h,p,a,l\}}}\left\{\lambda\frac{1}{\chi}Y^{1-\chi}\left[sn_y^{\theta}k_y^{\gamma}\right]^{\chi}+H\ln\left(\frac{h^{1-\zeta}b_r^{\zeta}}{p}\right)p-\phi\left(n_y+n_h\right)^{\pi}p^{1-\pi}\right.$$

$$-\lambda r_k\left(k_y+k_h\right)-\lambda\eta\left(a+\Lambda x-l\right)$$

$$\left.-\frac{A}{2}\left(\frac{a}{x}\right)^2x-\left[-\psi_1\frac{l}{x}+\frac{\psi_2}{2}\left(\frac{l}{x}\right)^2\right]x+\beta\int V\left(z'\right)dQ\left(s';s,s_{-1}\right)\right\}$$

subject to

$$\begin{aligned} p &= x+a+\Lambda x-l \qquad (14)\\ n_y+n_h &\leq p\\ b_r+b_h &= 1 \end{aligned}$$

plus (12), (13), and non-negativity constraints on the choice variables.

The planner's objective is to maximize the expected present discounted value of net local surplus. To see this note that the first two terms are the value of intermediate good production and the housing services consumed in the city. The next five terms comprise the contemporaneous costs to the planner of obtaining this surplus: the disutility of sending the indicated number of people to work; the shadow cost of equipment used in the city; and the disutility of net migration inclusive of guided trip production and location-taste shocks. The last term is the discounted continuation value given the updated state vector. Constraining the achievement of the city planner's objective are the local resource constraints, the housing and population transition equations and the non-negativity constraints on the choice variables. Note that in the statement of the land constraint we have normalized the local endowment of residential land to unity and used the fact that land used for current housing services cannot be built on in the same period.[17]

The unique stationary allocation is the solution to the city-planner's problem that satisfies particular side conditions. Let $\{n_y, n_h, k_y, k_h, h', b_r, b_h, p, a, l\}$ denote the optimal decision rules (which are functions of the state z) for the city planner's problem that takes $\{Y, \lambda, \eta, r_k, \Lambda\}$ as given and let μ be the invariant distribution generated by the optimal decision rules $\{h', p\}$ and the transition function Q. Define the aggregate stock of equipment, $K = \int (k_y + k_h)\, d\mu$, and per capita consumption, $C = Y - \delta_k K$, where $0 < \delta_k < 1$ denotes equipment's depreciation rate. Now suppose the following equations are satisfied

$$Y = \left\{ \int \left[s n_y (z)^{\theta} k_y (z)^{\gamma} \right]^{\chi} d\mu \right\}^{\frac{1}{\chi}} \tag{15}$$

$$\lambda = \frac{1}{C} \tag{16}$$

$$\int a(z)\, d\mu + \Lambda = \int l(z)\, d\mu \tag{17}$$

$$r_k = \frac{1}{\beta} - 1 + \delta_k \tag{18}$$

$$\lambda \int [\xi(z) - \eta]\, x d\mu - \tau \leq 0, (= 0 \text{ if } \Lambda > 0) \tag{19}$$

Then $\{C, K, n_y, n_h, k_y, k_h, h', b_r, b_h, p, \Lambda, a, l\}$ is a steady state allocation.[18] Equation (15) expresses aggregate output in terms of each city's intermediate good production. This equation is the theoretical counterpart to equation (2) used to estimate city-specific TFP. The marginal utility of consumption is given by equation (16). Equation (17) states that total in-migration equals total out-migration.

[17] In our calibration $0 < \theta + \gamma < 1$ which implies there is a fixed factor in intermediate good production that is constant (equal to one) across cities. One interpretation of this fixed factor is that it represents commercial land. Under this interpretation commercial land cannot be converted into residential land and vice versa.

[18] See the online technical appendix for a proof of this claim. We describe how we solve the model in Appendix F.

Equation (18) defines the rental rate for equipment. The last side condition (19) determines steady state undirected migration. It turns out to be identical to the first order condition for Λ in the static model.

The function $\xi(z)$ in (19) represents the value to the city planner of bringing an additional individual to the city. It is integral to migration's determination in the model and can be shown to satisfy

$$\xi(z) = C\phi\left[n_y(z) + n_h(z)\right]^{\pi}(\pi - 1)\, p(z)^{-\pi} + CH \ln\left(\frac{h(z)^{\varsigma}\, b_r(z)^{1-\varsigma}}{p(z)}\right) - CH$$
$$+ \beta \int \left(CA\left[\frac{a(z')}{p(z)}\right]^2 + C\psi_2\left[\frac{l(z')}{p(z)}\right]^2 + \Lambda\left[\xi(z') - \eta\right] + \xi(z')\right) dQ(s'; s, s_{-1}). \tag{20}$$

The value of bringing an additional worker to a city comprises four terms: the benefits of obtaining a better selection of worker disutilities given the same amount of total employment $n_y + n_h$; the benefits of the local housing services that the additional person will enjoy; the costs of reducing housing services for everybody else; and the expected discounted value of starting the following period with an additional person. This last term includes the benefits of having an additional person producing guided trips to the city, the benefits of obtaining a better selection of location-taste shocks, and the benefits of attracting additional people to the city through undirected migration.

If $a(z) = l(z) = 0$ never occurs in equilibrium, then using the first order conditions of the city-planner's problem one can show that

$$\lambda\xi(z) = \begin{cases} A\left[\frac{a(z)}{x}\right] + \lambda\eta, \text{ if } a(z) > 0, \\ \left[\psi_1 - \psi_2\left(\frac{l(z)}{x}\right)\right] + \lambda\eta, \text{ if } l(z) > 0. \end{cases} \tag{21}$$

Comparing equation (21) to the first order conditions for $a(z)$ and $l(z)$ in Section

3's static model we see that if gross migration rates are positive then the shadow value of a migrant is related to migration costs as in the static model. When there are no migration frictions, $A = \psi_1 = \psi_2 = 0$, equation (21) implies that the marginal value of bringing an additional individual to a city is equated across cities, $\xi(z) = \eta$, $\forall z$. However, this does not imply that wages are equated across cities. Instead, equation (20) says that the marginal savings in worker disutility plus the marginal impact on the utility of housing services is equated. When in addition to $A = \psi_1 = \psi_2 = 0$ housing structures are made perfectly mobile across cities, the same condition is obtained because land remains immobile. Finally, when land is also made mobile, then the marginal savings in work disutility and the marginal utility of housing services are each equated across cities.

5. Calibration

We now calibrate the steady state competitive equilibrium to U.S. data.[19] Our calibration has two important characteristics. First, the city-specific TFP process is chosen to match the estimates presented in Section 2.2 thereby pinning down the model's exogenous source of persistence and volatility. 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 TFP shock. The model's response of population to a TFP shock is the consequence of the estimated TFP process and the remaining parameters that are chosen to fit other features of the data.

In addition to specifying the stochastic process for TFP we need to find values for 16 parameters:

$$\theta, \gamma, \alpha, \vartheta, \delta_k, \delta_h, \beta, H, \zeta, \pi, \phi, \psi_1, \psi_2, A, \tau, \chi.$$

[19]Except where noted the aggregate data used to calibrate our model is obtained from Haver Analytics.

These include the factor shares in production and construction, depreciation rates for equipment and structures, the discount factor, the housing coefficient in preferences, land's share in housing services, and the parameters governing labor supply, migration, and intermediate goods' substitutability.

We calibrate these parameters conditional on a given quantity of undirected migration Λ determined by τ. For larger values of τ undirected migration is relatively small so that $a(z) > 0,\ \forall z$. In these cases the behavior of the model is invariant to the specific value of τ. For smaller values of τ undirected migration is large and $a(z) = 0$ for some z. In these cases the behavior of the model is affected. It turns out that even for seemingly large steady state Λ corner solutions for $a(z)$ are either non-existent or extremely rare. We set our baseline so that the undirected arrival rate is 3.8%, roughly 70% of all moves.[20]

The baseline calibration for the assumed value of τ is summarized in Table 1. There we indicate for each parameter the proximate calibration target, the actual value for the target we obtain in the baseline calibration, and the resulting parameter value. In the remainder of this section we discuss the calculations underlying Table 1. We begin with the novel aspects of our calibration which involve the parameters governing migration, the city-level TFP process, the elasticity of substitution of intermediate goods, and labor supply.

5.1. Migration Parameters

Section 3.3 establishes that the migration parameters A, ψ_1 and ψ_2 are essential for determining the speed of population's adjustment to TFP shocks in the model. Fortunately there is evidence at hand that makes assigning values to these parameters straightforward. First, conditional on a value for A reproducing Figure

[20]The specific value is $\tau = 1$. For this value the baseline calibration has 0.3% of city-year observations involving zero arrivals.

Table 1: Baseline calibration

Parameter	Parameter Description	Calibration Target	Target Value	Actual Value	Parameter Value
θ	Labor's share in intermediate goods	$\int w\left[n_y+n_h\right]d\mu/GDP$	0.64	0.64	0.66
α	Labor's share in construction	$\int n_h d\mu/\int\left[n_y+n_h\right]d\mu$	0.042	0.042	0.41
β	Discount factor	Real interest rate	0.04	0.04	0.9615
γ	Intermediate goods' equipment share	K_y/GDP	1.53	1.53	0.235
δ_k	Depreciation rate of equipment	$\delta_k K/GDP$	0.16	0.16	0.104
ϑ	Equipment's share in construction	K_h/GDP	0.022	0.022	0.05
δ_h	Depreciation rate of structures	I/GDP	0.064	0.064	0.045
ζ	Land's share in housing services	$\int q^b b_r d\mu/\left[\int q^h h d\mu+\int q^b b_r d\mu\right]$	0.37	0.36	0.215
H	Housing coefficient in preferences	$\int q^h h d\mu/GDP$	1.55	1.50	0.205
ϕ	Labor disutility	$\int\left[n_y+n_h\right]d\mu/\int p d\mu$	0.63	0.63	1.61
π	Determines labor supply elasticity	$\partial\ln\left[n_y+n_h/p\right]/\partial\ln w$	0.24	0.25	5.0
ψ_1	Taste shock slope	Mean arrivals	5.5	5.5	6.12
ψ_2	Taste shock curvature	Slope of arrivals versus net	0.57	0.57	44.4
A	Guided trip cost	Average moving costs/average wages	-1.9	-1.9	33.5
χ	Intermediate goods' complimentarity	Zipf's law for population	-1.0	-1.3	0.9
g	Drift in technology	Zipf's law for TFP	-3.5	-3.4	-0.0017
ρ	TFP lag coefficient	Serial corr. of TFP growth	0.24	0.22	0.35
σ	TFP innovation std. err.	TFP growth innovation std. err.	0.015	0.015	0.019

Note: The calibration is based on $\tau=1$. See the text for details. The actual values for the serial correlation and innovation of TFP growth are based on simulations. The underlying parameter values for the TFP process are somewhat different because of the discrete approximation we use.

2 pins down ψ_1 and ψ_2 via equation (10).[21] In particular

$$\frac{\psi_1}{A+\psi_2} + \frac{A}{A+\psi_2}\Lambda = 5.5 \text{ and } \frac{\psi_2}{A+\psi_2} = 0.57.$$

Using the slope of the gross versus net migration relationship in Figure 2 to measure the migration cost parameters is similar to Gavazza, Mongey, and Violante (2014) who identify parameters in a hiring cost function by relating them to the slope of the log-linear relationship between the job filling rate and employment growth at the firm level estimated by Davis, Faberman, and Haltiwanger (2013).

To identify A we take advantage of Kennan and Walker (2011)'s estimate of the average net cost of migration for those who move. Specifically, we match the statistic defined as the average net cost of migration of those who move divided by average wages where we take the latter from Kennan and Walker (2011) as well.[22] It is straightforward to replicate their concept of moving costs in our model. In Kennan and Walker (2011), net moving costs sum two components of the utility flow of an individual in the period of a move. One component called "deterministic moving costs" is a function of the distance of the move, whether the move is to a location previously visited or not, the age of the mover, and the size of the destination location. The second component is the difference between idiosyncratic benefits in the current and destination location. We interpret the cost of guided trips and undirected migration as representing the first component and the location-taste shocks the second one. Consequently we measure average

[21]The constant term is the sample average gross migration rate. Calibrating to the average the start or end of our sample leads to very similar results.

[22]Using Kennan and Walker (2011)'s estimates, average net moving costs of those who move divided by average annual wages equal -1.9. This value equals the ratio -\$80,768/\$42,850. The numerator is the entry in the row and columns titled 'Total' in Table V and the denominator is the wage income of the median AFQT scorer aged 30 in 1989 reported in Table III. The negative value of the estimate indicates that individuals receive benefits to induce them to move.

moving costs of individuals who move as

$$\frac{\int q(z)a(z)\ d\mu + C\tau\Lambda}{\int a(z)\ d\mu + \Lambda} + \frac{C\int\left(-\psi_1\frac{l(z)}{x} + \psi_2\left(\frac{l(z)}{x}\right)^2\right)xd\mu}{\int l(z)d\mu},$$

where $q(z)$ is the price of a guided trip in the decentralization discussed in Appendix C. Average wages are simply

$$\frac{\int w(z)\left[n_y(z) + n_h(z)\right]d\mu}{\int\left[n_y(z) + n_h(z)\right]d\mu},$$

where wages in a type-z city, $w(z)$, equal the marginal product of labor.

There are two potential drawbacks to using Kennan and Walker (2011)'s estimate of moving costs. First, they identify moving costs using individual-level data. Since an individual's location choice is not determined in the model we cannot replicate their estimation strategy. Still, our model implies a value for Kennan and Walker (2011)'s moving cost statistic so it is natural to take advantage of their estimates. Of more concern is that Kennan and Walker (2011) estimate moving costs using data on the frequency of inter-state moves, while our quantitative model describes inter-city moves. Inter-city moves are more frequent than moves between states. Consequently it is possible that Kennan and Walker (2011) would have estimated a different value for moving costs had they had data on all inter-city moves, in which case we would be calibrating our model to the wrong value. This suggests it is important to quantify the bias in Kennan and Walker (2011)'s estimate. In Appendix E we study a calibrated variant of their model and find that the bias is likely to be small.

5.2. TFP and Substitutability of City-specific Goods

The calibration of the substitution parameter χ and the stochastic process for TFP are interconnected because χ is used to measure TFP. When we measure TFP using the procedure described in Section 2.2 its growth rate is well-represented as a stationary AR(1) process which is non-stationary in levels and therefore inconsistent with a steady state. To overcome this we assume a reflecting barrier process for TFP:

$$\ln s_{t+1} = \max \left\{ g + (1+\rho) \ln s_t - \rho \ln s_{t-1} + \varepsilon_{t+1}, \ln s_{\min} \right\}. \tag{22}$$

where $\varepsilon_{t+1} \sim N(0, \sigma^2)$, $g < 0$ and $\rho > 0$. With this process TFP growth is approximately AR(1), while its level is stationary due to having a negative drift and being reflected at the barrier $\ln s_{\min}$ (normalized to zero). The model's endogenous variables will appear to be non-stationary over samples of similar length to our data.

Gabaix (1999) assumed $\rho = 0$ to explain the cross section distribution of cities by population. In this case the invariant distribution has an exponential upper tail given by

$$\Pr \left[s_t > b \right] = \frac{d}{b^{\omega}}$$

for scalar parameters d and ω. A striking characteristic of cities is that when s measures a city's population one typically finds that $\omega \simeq 1$. Equivalently a regression of log rank on log level of city populations yields a coefficient close to -1. This property is called Zipf's law and so we refer to ω as the Zipf coefficient. The case $\rho > 0$, which applies when TFP growth is serially correlated, has not been studied before. Simulations suggest this case also has an invariant distribution with an exponential-like upper tail. We verify below that a version of Zipf's law holds for TFP and so using the reflecting barrier process with $\rho > 0$ seems justified.

Our calibration of χ and (22) proceeds as follows. For a given χ (and θ and γ which are calibrated independently as discussed below) we measure TFP in the data following the procedure in Section 2.2, obtain its Zipf coefficient, and estimate an AR(1) in its growth rate. We then find the g, ρ and σ to match the Zipf coefficient and the parameters of the estimated AR(1) using data simulated from our model and based on all of these parameters calculate the model's population Zipf coefficient. The calibrated value of χ is the one that generates a population Zipf coefficient that is as close as possible to the one we find in the data, 1.0. The best fit is at $\chi = 0.9$ with a population Zipf coefficient equal to 1.3. The corresponding values of g, ρ and σ are shown in Table 1.

Figure 5: Zipf's laws for population and TFP

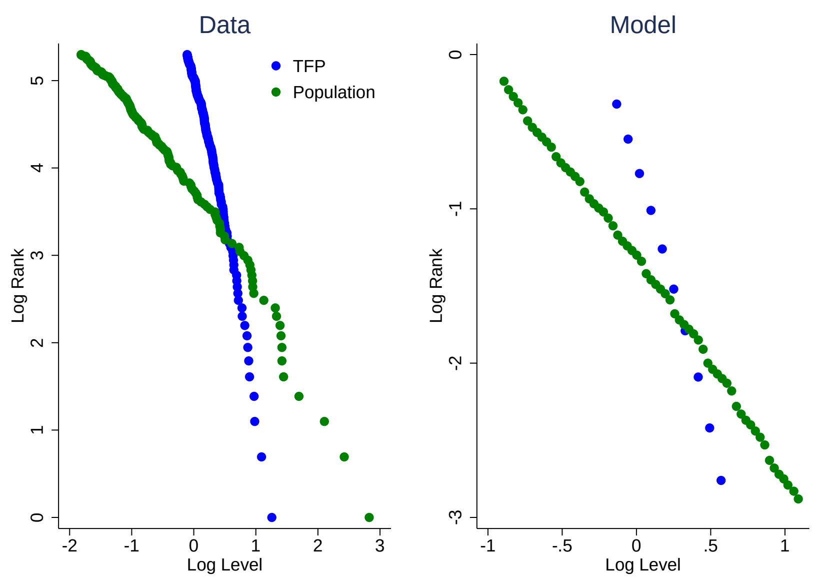


Figure 5 demonstrates the model's success at replicating the two Zipf's laws by plotting log rank versus log level for population and TFP with empirical and simulated data.[23] Log level of TFP is calculated using (4) without the first

[23]The left plot includes the top 200 cities by population in 1990. The right plot excludes the lower 5% of cities to be comparable, and the top 5% due to the finite upper bound of the TFP domain which introduces bunching in the right tail of the population distribution. We measure rank in the model with cumulative distribution functions. TFP's domain is narrower in the model to reduce computational costs. The narrower domain does not matter for our calibration. For example, the migration parameters are based on Figure 2 and Kennan and Walker (2011)'s

difference operator. Notice that TFP's Zipf coefficient is larger than population's in the data. This feature arises naturally in the model because population tends to be allocated away from lower toward higher TFP cities. Equivalently, the long run response of population to a TFP shock is larger than that of TFP. Luttmer (2007) finds a similar relationship between employment and TFP in an equilibrium model of firm size.

5.3. Labor Supply

Labor supply is determined by ϕ and π. The former is chosen to match the ratio of civilian employment to population obtained from Census Bureau data and the latter is chosen using the first order condition for labor supply. In the model's decentralization the representative household chooses labor supply in a city to equate the disutility of working an additional household member to the local wage. This implies:

$$(1-\pi)\left(\Delta\hat{n}_{it} - \Delta\hat{p}_{it}\right) + \Delta\hat{w}_{it} = 0,$$

where $n = n_y + n_h$. Using the methods described in Section 2.2, we estimate the dynamic responses of $\Delta\hat{n}_{it}$, $\Delta\hat{p}_{it}$ and $\Delta\hat{w}_{it}$ to a local TFP shock and calibrate π so that this equation holds in the period of a shock. This procedure places no restrictions on the *levels* of the variables' impulse responses. The resulting Frisch elasticity, .25, is surprisingly close to micro estimates of the extensive margin elasticity discussed by Chetty, Guren, Manoli, and Weber (2011).

estimate of migration costs. In the model the former does not depend on the level of TFP and the latter depends on the distribution of TFP growth which is essentially independent of the domain. Our quantitative analysis is based on growth rates and so is similarly independent of the underlying domain.

5.4. Remaining Parameters

The remaining parameters are calibrated similarly to studies based on the neoclassical growth model. Several targets involve GDP and we measure this in the model as

$$GDP = Y + I, \tag{23}$$

where Y is output of non-construction final goods and I is residential investment. Residential investment is measured as the value in contemporaneous consumption units of the total additions to local housing in a year. Specifically,

$$I = \int \left[\beta \int q_h(z')dQ(s'; s, s_{-1}) \right] n_h(z)^{\alpha} k_h(z)^{\vartheta} b_h(z)^{1-\alpha-\vartheta} d\mu$$

where q_h denotes the price of residential structures. This price is obtained as the solution to the following no arbitrage condition

$$q_h(z) = r_h(z) + (1 - \delta_h) \beta \int q_h(z') \, dQ(s'; s, s_{-1})$$

where the rental price of residential structures, r_h, equals the marginal product of structures in the provision of housing services. The National Income and Product Accounts (NIPA) measure of private residential investment is the empirical counterpart to I. Our empirical measure of Y is the sum of personal consumption expenditures less housing services, equipment investment and private business inventory investment. Because our model does not include non-residential structures investment, government expenditures and net exports we exclude these from our empirical concept of GDP.

Our measurement of model GDP and wages excludes the value of guided trip services, which might be problematic. For example, workers produce guided trips and in principle they should be compensated for this. We calculate the total value of guided trips in our baseline calibration to be 1.8% of model GDP.

The decentralization described in Appendix C suggests we can interpret guided trips as encompassing many market and non-market activities. Some of these activities appear in the national accounts as business services and therefore count as intermediate inputs that appear only indirectly in measured GDP. Others do not appear anywhere in the national accounts because they are essentially home production or are impossible to measure. Fortunately, given its small size, including the total value of guided trips in model GDP and wages does not change our baseline calibration.

Measuring employment also is complicated by the fact that all household members participate in generating guided trips. We count those agents engaged in intermediate good production, n_y, and residential construction, n_h, as employed and measure their wages by their marginal products excluding the value of guided trips. The non-employed who also produce guided trips are assumed to be engaged in home production and so are not included in our accounting of employment. In Table 1 the labor share parameters are chosen to match total labor compensation as a share of GDP (the target is borrowed from traditional real business cycle studies) and our estimate of the share of residential construction employment in total private non farm employment.[24]

We fix the discount rate so the model's real interest rate is 4%. Combined with this target the equipment-output ratio in the non-construction sector, K_y/Y, identifies equipment's share in that sector's production. Our empirical measure of equipment for this calculation is the Bureau of Economic Analysis' (BEA) measure of the stock of equipment capital. Equipment's depreciation rate is identified using the investment to GDP ratio, where we measure investment using the NIPA estimate of equipment investment. Equipment's share in residential construction

[24]We estimate housing construction employment by multiplying total construction employment by the average over our sample period of the nominal share of residential investment in total structures investment expenditures.

is identified by the ratio of capital employed in the residential construction sector, K_h, to GDP where the empirical counterpart to capital in this ratio is the BEA measure of equipment employed in residential construction. The depreciation rate of residential structures is identified using the residential investment to GDP ratio.[25]

We identify the housing service parameters as follows. First the housing coefficient H is chosen to match the residential capital to GDP ratio, where the measurement of residential capital is consistent with our measure of residential investment described above. Land's share in housing services, ζ, is chosen to match the estimate of land's share of the total value of housing in Davis and Heathcote (2007). To measure this object in the model we need the price of land, q_b. We obtain this variable as the solution to the no-arbitrage condition

$$q_b(z) = r_b(z) + \beta \int q_b(z')\, dQ(s'; s, s_{-1}),$$

where r_b denotes the rental price of land which equals the marginal product of land in the provision of housing services. Land's share of the economy-wide value of housing is then given by $\int q_b b_r d\mu / \left[\int q_h h d\mu + \int q_b b_r d\mu\right]$.

6. Quantitative Analysis

We now consider the model's empirical predictions. First, we confirm that the model accounts for population's slow adjustment to TFP shocks and that this success comes with generally accurate predictions for the other variables we study. Next, we assess the model's unconditional dynamics and find it is similarly successful even though TFP shocks are the only source of fluctuations in the model.

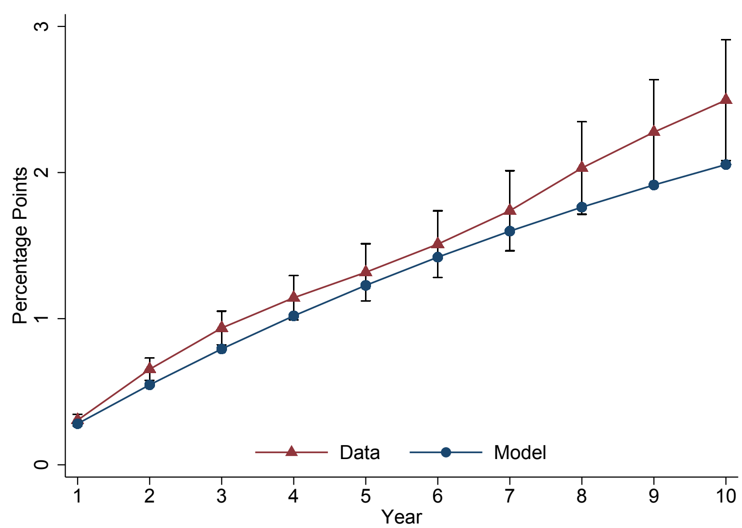
[25] The depreciation rate for residential structures obtained this way is close to the mean value of the depreciation-stock ratio for housing obtained from the BEA publication "Fixed Assets and Consumer Durable Goods," once output and population growth are taken into account. Calibrating to this alternative value has virtually no impact on our quantitative findings.

After establishing the empirical plausibility of our framework, we investigate how migration and housing influence slow population adjustments.

6.1. Responses to Productivity Shocks

We estimate the responses to TFP shocks of population, gross migration, employment, wages, residential investment and house prices in both the model and the data using the procedure described in Section 2.2, basing our model responses on the simulation of a large panel of cities over a long time period. TFP is measured the same way with model generated and empirical data.

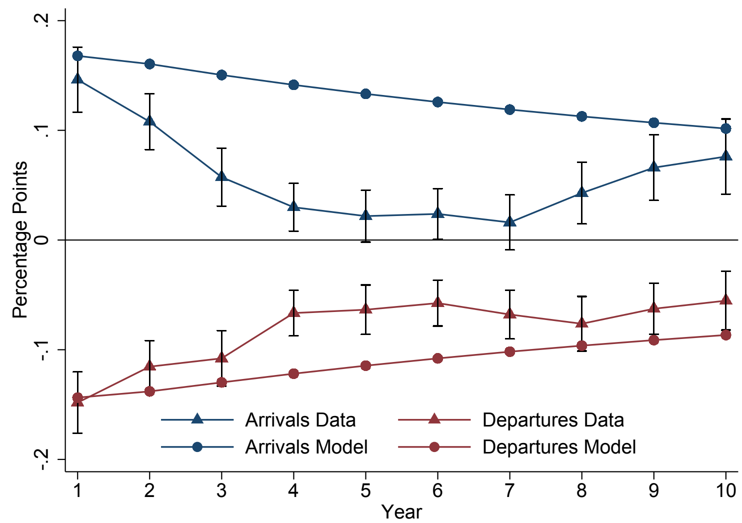
Figure 6: Response of population



Note: Point estimates along with 2 standard error bands.

Figure 6 displays model and estimated responses of population to a one standard deviation positive innovation to TFP. It demonstrates that the model's population response is statistically and economically close to the one for the data. In other words our model accounts for the slow response of city populations to local TFP shocks. Note that the model is able to account for the unconditional population distribution depicted in Figure 5 with a slow conditional response of

Figure 7: Responses of arrival and departure Rates



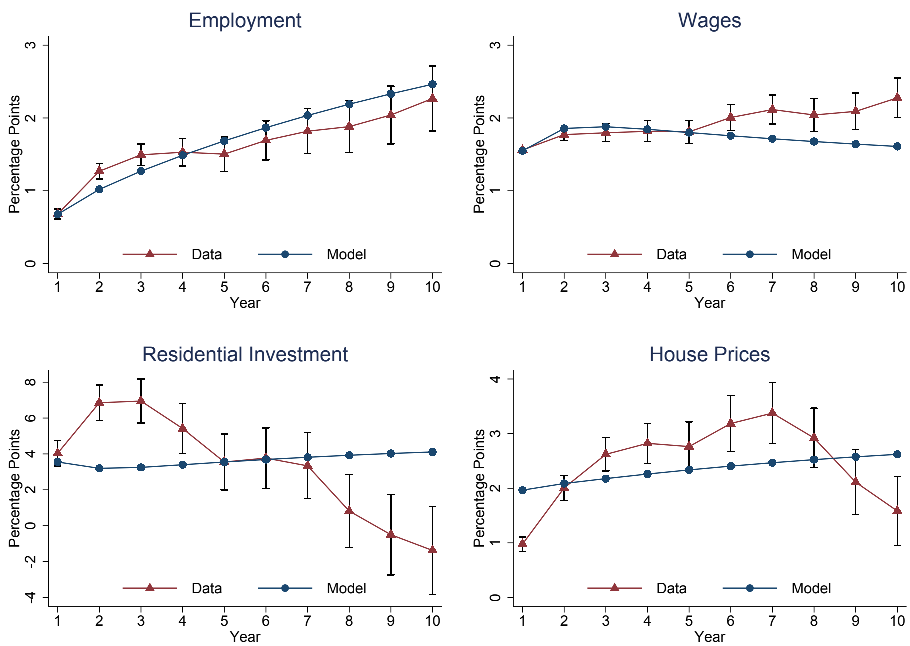
Note: Point estimates along with 2 standard error bands.

population to a TFP shock because ultimately the long run response of population exceeds that for TFP.

Figure 7 shows this accounting for slow population adjustments involves replicating the dynamic responses of the arrival and departure rates quite well. Crucially the model is consistent with the negative conditional correlation between the gross migration rates. The intuition for this finding is simple. Having multiple margins to respond to the increase in productivity, the city planner takes advantage of all of them. It can raise employment per person and bring more workers to the city. For the latter it can cut back on the fraction of the initial population that leaves for other cities and attract more workers with guided trips. The goodness-of-fit is weaker for gross migration than it is for population. Nevertheless given its simplicity the model does surprisingly well.

Figure 8 shows the dynamic responses of employment, wages, residential investment and house prices. We define house prices, q_{sf}, as the total value of structures and land used to produce housing services per unit of housing services

Figure 8: Responses of labor and housing markets



Note: Point estimates along with 2 standard error bands.

provided:

$$q_{sf}\left(z\right)=\frac{q_{h}\left(z\right)h\left(z\right)+q_{b}\left(z\right)b_{n}(z)}{h(z)^{1-\varsigma}b_{r}(z)^{\varsigma}}.$$

The price q_{sf} corresponds to the price of housing per square foot under the assumption that every square foot of built housing yields the same quantity of housing services.

The labor market responses are a very good fit. Observe that in both the model and the data the initial employment response is stronger than population's. That is, the employment to population ratio rises after a positive TFP shock indicating that the labor supply margin is indeed exploited in both the data and the model. The responses of construction and housing are broadly consistent with the data with the orders of magnitude about right. These findings derive from a higher population desiring additional housing and an imperfectly elastic supply of new housing due to competing uses for the local labor needed to produce it.[26]

6.2. Unconditional Statistics

Our estimates of responses to TFP shocks are robust to the presence of other shocks which are likely to be present in the data, for example shocks to local taxes and amenities. However, if our model accounts for only a small fraction of the total variation in the data we would be less confident in its predictions. Consequently in Table 2 we display standard deviations, contemporaneous correlations, and auto-correlations at the annual frequency in the model and in our data. Standard deviations of all variables except population are expressed relative to the standard deviation of population and all contemporaneous correlations are with population.

Table 2: Volatility, co-movement and auto-correlation within cities

	Standard Deviation		Correlations		Auto-Correlations	
Variable	Data	Model	Data	Model	Data	Model
Population	1.35	0.87	–	–	0.81	0.93
Arrival Rate	0.67	0.53	0.59	1.00	0.81	0.93
Departure rate	0.58	0.48	-0.47	-1.00	0.81	0.93
Employment	1.54	1.23	0.57	0.93	0.51	0.73
Wages	1.22	1.81	0.16	0.32	0.14	0.20
Res. Investment	20.5	4.27	0.09	0.40	-0.04	0.09
House Prices	3.03	2.32	0.28	0.47	0.69	0.07

Note: The statistics are based levels of the gross migration rates and on the growth rates of the other variables. All variables have been transformed as described in Section 2. Standard deviations of all variables except population are expressed relative to the standard deviation for population. Correlations are with population.

Table 2 indicates that TFP shocks are a significant source of local variation: the standard deviation of annual population growth in the model is about two

[26] Lloyd-Ellis et al. (2014) demonstrate that search frictions show promise in generating serially correlated responses of construction and house price growth to productivity shocks.

thirds of that in the data. The model is strikingly successful at replicating the qualitative pattern of relative volatilities and, with the exception of residential construction, the magnitudes are about right as well. Correlations with population growth are also similar in the model and the data.[27] Finally, with the exception of house prices, persistence in the model corresponds well with the data. Notice in particular that the gross migration rates and population growth are identically auto-correlated in the data, just as they are in the model. This correspondence reflects the linear relationship between gross and net migration we have uncovered in the data and have built into our model.

While Table 2 displays persistence at the annual frequency much of the literature focuses on longer run persistence. For example Blanchard and Katz (1992), Glaeser et al. (1995) and Rappaport (2004) report highly persistent 10-year population and employment growth. In our data the auto-correlations of 10-year population and employment growth are .88 and .65, respectively. In our model the corresponding correlations are .63 and .58. Evidently our model generates almost as much long-run persistence in population and employment growth as we find in the data. This is important for understanding our model's success at accounting for persistent urban decline discussed below.

6.3. The Source of Slow Population Adjustments

We now disentangle the sources of slow population adjustments in our model. To do so we consider several perturbations of the baseline calibration. In the *immobile housing only* case we fix $A = 0$. When $A = 0$ the departure rate in each city equals the constant value that minimizes departure costs and population adjusts via costless changes in the arrival rate. Therefore in this case only the

[27] The largest discrepancies involve the arrival and departure rates being perfectly positively and negatively correlated with population growth. Including a mechanism to reproduce Kennan and Walker (2011)'s finding that out-migration is relatively high for recent in-migrants might move the model in the direction of the data.

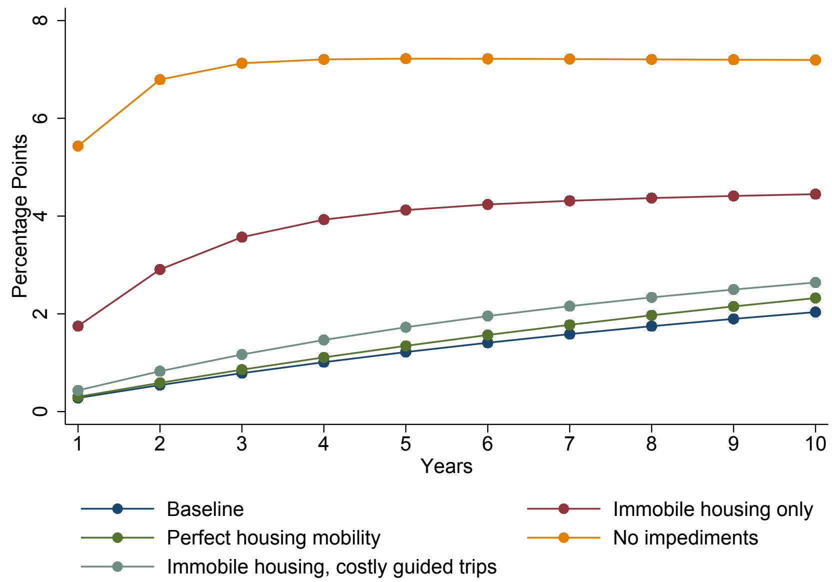
factors inducing inelastic local housing supply impede population's response to a TFP shock. In the *immobile housing, costly guided trips* case we retain housing's immobility and set $\psi_1 = \psi_2 = 0$ so that costly guided trips are the only migration-related impediment to population adjustments. Under *perfect housing mobility* the local supply of housing is perfectly elastic at the economy-wide rental rate. In this case only the migration costs impede population adjustments. The *no impediments* case combines all of these perturbations so there are no migration costs and housing is perfectly mobile.

Figure 9 displays population's response to TFP shocks over 10 years for these four perturbations along with the response implied by the baseline calibration. In the no impediments case population's initial response is almost 20 times the baseline and it converges quickly to its new long run level, closely tracking TFP.[28] Over an infinite horizon the response of population is pinned down by the size of the TFP shock – in every perturbation population converges to the level it reaches when there are no impediments to its adjustment. However it can take a very long time to get there.

The population response under immobile housing only is substantially slower than in the no impediments case, but it is far from the baseline and hence the data as well. The discrepancy between the no impediments and the immobile housing only cases arises from a property of adjustment costs highlighted by Abel and Eberly (1994). They demonstrate that the first adjustment cost introduced to an otherwise frictionless model always has a relatively large impact on dynamics compared to when additional costs are introduced. The immobile housing, costly guided trips case is quite close to the very slow baseline response: the costs of

[28]In the no impediments case aggregate consumption is about 5% higher than in our baseline economy, while employment is about the same. This suggests that the combination of migration costs and housing immobility can generate a substantial amount of labor mismatch in the economy. See Sahin, Song, Topa, and Violante (2014) for an analysis of mismatch due to search frictions in the labor market.

Figure 9: Why is population slow to adjust to shocks?



finding new cities to live and work have large effects even when housing is already immobile.

These two perturbations suggest the slow baseline response requires both adjustment costs in housing and migration. However notice that the perfect housing mobility case is very close to the baseline case. Essentially it is not necessary to introduced inelastic local housing supply to get close to the data. This finding suggests that migration costs are the prime driver of slow population adjustments. Given the large impact of introducing costly guided trips when housing is already immobile we infer that the costs associated with finding new cities to live and work are the most important impediment to population adjustments.

The finding of slow population adjustments driven mostly by costly migration confirms and reinforces results in Kennan and Walker (2011). Using the parameters of a migration choice problem estimated with data on the frequency of inter-state moves, they calculate the optimal responses of individuals to a permanent wage change in one state. From these choices they obtain a matrix of transition probabilities from which they trace out the response of population in the state with the permanent wage change. Strikingly, we find roughly the same five-year elasticity of population with respect to the wage, about 0.5.[29] In Kennan and Walker (2011) about 30% of the five year response occurs in the first year (see their Figure 1), whereas we find a response closer to 20%. The slower response we obtain comes from wages taking several periods to reach their new long run level compared to the immediate change in Kennan and Walker (2011). Our slower wage response derives from the endogenous response of wages induced by net in-migration. Figure 9 suggests that inelastic local housing supply is not an important source of the difference. Overall our results establish that Kennan and Walker (2011)'s findings are largely robust to the presence of inelastic local housing supply and equilibrium interactions as well as considering migration between cities instead of states.

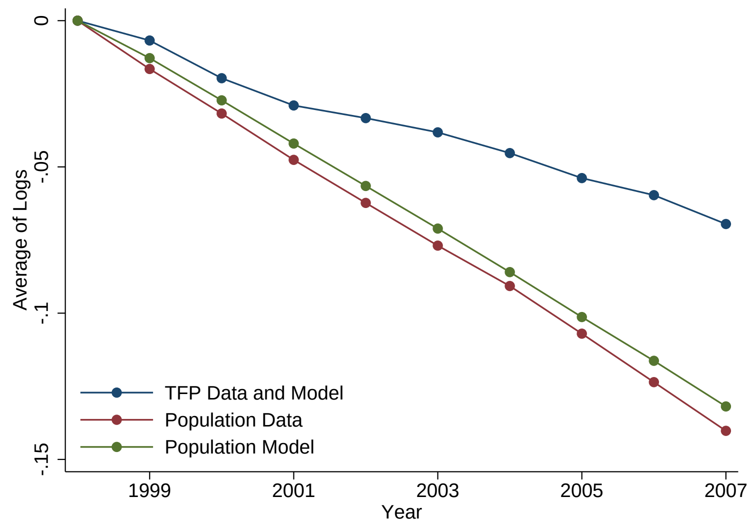
6.4. Migration and Urban Decline

There are many cities which experience declining populations over the sample period 1985-2007. This is evidence of the persistent urban decline studied by Glaeser and Gyourko (2005). Interestingly, the cities with declining populations also have relative TFP declining for most of the sample. The slow response of population to TFP shocks in our model then suggests it might be able to account for persistent urban decline. We now discuss a simple experiment to demonstrate that indeed our calibrated model does account for persistent urban decline.

We focus on the 15 cities that experience the greatest population declines in our sample. The corresponding TFP paths are fed into the model from the common initial condition that takes the mid-point of our TFP grid and assumes TFP stays at that level for a long time. Then we use the first 12 years of our

[29] Kennan and Walker (2011) consider a 10% increase in wages and find that population is 5% higher after five years. We find a 1.1% response of population to a 2% permanent increase in the wage.

Figure 10: Persistent urban decline



sample, 1985 to 1997, to simulate unique initial conditions for each city based on its empirical TFP path. This procedure builds in the possibility that past declines in TFP show through into future population declines. For each city we calculate the predicted path for log population starting in 1998, average over these paths and compare the result to the same object constructed using the data.[30]

Figure 10 shows the average log paths for TFP and population for the data and the model. TFP falls by .06 from 1998-2007 and population falls by twice as much. The model's predicted path for population lies very close to its empirical counterpart. Obviously the fit is not as perfect for the individual cities, but the general impression is similar. There are two key factors driving the model's success: persistent declines in TFP taken from the data and the very slow response of population to past declines in TFP predicted by the model. The impact of past TFP declines on current population is demonstrated by the faster rate of population decline relative to TFP – in the short run population's response to a

[30]For this experiment we equate our empirical measure of TFP to model TFP, s. We do this for computational tractability but the small differences involved should not affect our conclusions.

TFP innovation is much smaller than TFP's, but over longer horizons it responds by much more.

Since the predominant source of slow population adjustment in the model is the cost of finding new cities, we conclude that these costs are integral to our model's explanation of persistent urban decline. Durable and immobile housing is not important at all in the sense that the intrinsic costs of migration essentially account for slow population adjustments on their own. This contrasts with Glaeser and Gyourko (2005) who argue that durable and immobile housing explains persistent urban decline.[31] These authors do not integrate costly migration into their empirical analysis. We consider both housing and migration costs in a unified framework, but housing turns out to be relatively unimportant.

7. Conclusion

This paper documents that a city's population adjusts slowly to near random-walk TFP shocks and shows that empirically plausible inelastic local housing supply and migration costs can explain why. We build a model that allows us to disentangle the effects of housing and migration on urban population dynamics and find that inelastic local housing supply plays a relatively limited role compared to the intrinsic costs of migration and of the migration costs we consider those associated with finding new cities in which to live and work are the most important. We also show that our model can account for the persistence of urban decline. Our modeling of migration and calibration of its intrinsic costs is not arbitrary, but is dictated by the nature of the relationship between gross and net migration that we uncover and the microeconomic evidence.

[31] They show that irreversible housing in declining cities has several empirical predictions which they verify. Our model does not share these predictions since the irreversibility constraint is never binding. This is due to the small variance of TFP innovations compared to housing's depreciation rate. The constraint does not bind in our data either as new building permits are always strictly positive. It presumably binds for neighborhoods within a city and this might explain Glaeser and Gyourko (2005)'s findings.

Our model has left out interesting features that are undoubtedly important for understanding the full range of adjustment to shocks within and across cities. For example we have abstracted from search frictions in local labor and housing markets. We think it would be interesting to add these features to our framework. Doing so would help disentangle the contributions to labor reallocation of traditional search from the migration frictions we introduce. For example, local housing and labor search frictions might play a role similar to the migration frictions in our framework. It remains to be determined whether these frictions can account for the empirical relationship between gross and net migration and the slow response of population to TFP shocks that we uncover.

We abstracted from secular decline in rates of gross migration as well. While our model is able to explain persistent urban decline as the slow response of population to declines in local TFP, it is possible that this mechanism interacts with others that have been proposed to explain the trend. Understanding whether such an interaction exists seems worth exploring in future research.

References

Abel, A. B. and J. C. Eberly (1994). A unified model of investment under uncertainty. *American Economic Review 84*(1), 1369–1384.

Alvarez, F. and R. Shimer (2011). Search and rest unemployment. *Econometrica 79*(1), 75–122.

Alvarez, F. and M. Veracierto (2012). Fixed-term employment contracts in an equilibrium search model. *Journal of Economic Theory 147*(5), 1725–1753.

Blanchard, O. J. and L. F. Katz (1992). Regional evolutions. *Brookings Papers on Economic Activity 1992*(1), 1–75.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011). Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins. *American Economic Review Papers and Proceedings 101*, 471–475.

Coen-Pirani, D. (2010). Understanding gross worker flows across US states. *Journal of Monetary Economics 57*, 769–784.

Davis, M., J. Fisher, and T. Whited (2014). Macroeconomic implications of agglomeration. *Econometrica 82*(2), 731–764.

Davis, M. and J. Heathcote (2007). The price and quantity of residential land in the United States. *Journal of Monetary Economics 54*(8), 2595–2620.

Davis, M. and F. Ortalo-Mangé (2011). Household expenditures, wages and rents. *Review of Economic Dynamics 14*(2), 248–261.

Davis, S., J. Faberman, and J. Haltiwanger (2006). The flow approach to labor markets: New data sources and micro-macro links. *Journal of Economic Perspectives 20*(3), 3–26.

Davis, S., J. Faberman, and J. Haltiwanger (2013). The establishment-level behavior of vacancies and hiring. *Quarterly Journal of Economics 128*(2), 581–622.

Ferreira, F., J. Gyourko, and J. Tracy (2012). Housing busts and household mobility: An update. *Economic Policy Review November*, 1–15.

Gabaix, X. (1999). Zipf's law for cities: An explanation. *Quarterly Journal of Economics 114*(3), 739–767.

Gavazza, A., S. Mongey, and G. Violante (2014). What shifts the Beverage Curve: recruiting intensity and finanical shocks. New York Univeristy manuscript.

Glaeser, E. L. and J. Gyourko (2005). Urban decline and durable housing. *Journal of Political Economy 113*(2), 345–375.

Glaeser, E. L., J. A. Scheinkman, and A. Shleifer (1995). Economic growth in a cross-section of cities. *Journal of Monetary Economics 36*(1), 117–143.

Greenwood, M. (1997). Internal migration in developed countries. In M. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economics*, Volume 1B. North Holland.

Gyourko, J., A. Saiz, and A. Summers (2008). A new measure of the local regulatory environment for housing markets. *Urban Studies 45*(3), 693–729.

Kaplan, G. and S. Schulhofer-Wohl (2015). Understanding the long-run decline in interstate migration. Minneapolis Fed Working Paper 697.

Karahan, F. and S. Rhee (2013). Geographical reallocation and unemployment during the great recession: The role of the housing bust. Federal Reserve Bank of New York Staff Report No. 605.

Karahan, F. and S. Rhee (2014). Population aging, migration spillovers, and the decline in interstate migration. Federal Reserve Bank of New York working paper.

Kennan, J. (2008). Average switching costs in dynamic logit models. University of Wisconsin-Madison manuscript.

Kennan, J. and J. R. Walker (2011). The effect of expected income on individual migration decisions. *Econometrica 79*(1), 211–251.

Lloyd-Ellis, H. and A. Head (2012). Housing liquidity, mobility, and the labour market. *Review of Economic Studies 79*(4), 1559–1589.

Lloyd-Ellis, H., A. Head, and H. Sun (2014). Search, liquidity, and the dynamics of house prices and construction. *American Economic Review 104*(4), 1172–1210.

Lucas, R. (1997). Internal migration in developing countries. In M. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economics*, Volume 1B. North Holland.

Luttmer, E. (2007). Selection, growth, and the size distribution of firms. *Quarterly Journal of Economics 122*(3), 1103–1144.

Lyche, T. and K. Morken (2011). Spline methods. Department of Informatics, Centre of Mathematics for Applications, University of Oslo, manuscript.

Mangum, K. (2015). Cities and labor market dynamics. Georgia State University manuscript.

Modestino, A. S. and J. Dennett (2013). Are American homeowners locked into their houses? the impact of housing market conditions on state-to-state migration. *Regional Science and Urban Economics 43*(2), 322–337.

Molloy, R., C. L. Smith, and A. Wozniak (2011). Internal migration in the united states. *Journal of Economic Perspectives 25*(2), 1–42.

Nenov, P. (2015). Regional reallocation and housing markets in an equilibrium model of frictional migration. *Review of Economic Dynamics 18*(4), 863–880.

Rappaport, J. (2004). Why are population flows so persistent? *Journal of Urban Economics 56*, 554–580.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy 90*(6), 1257–1278.

Rosen, S. (1979). Wage based indexes of urban quality of life. In P. N. Miezkowski and M. R. Straszheim (Eds.), *Current Issues in Urban Economics*, pp. 74–104. Baltimore, MD: Johns Hopkins University Press.

Sahin, A., J. Song, G. Topa, and G. Violante (2014). Measuring mismatch in the US labor market. *American Economic Review 104*(11), 3529–3564.

Saiz, A. (2010). The geographic determinants of housing supply. *Quarterly Journal of Economics 125*(3), 1253–1296.

Saks, R. E. and A. Wozniak (2011). Labor reallocation over the business cycle: New evidence from internal migration. *Journal of Labor Economics 29*(4), 697–739.

Schulhofer-Wohl, S. (2012). Negative equity does not reduce homeowners' mobility. *Federal Reserve Bank of Minneapolis Quarterly Review 32*(1), 2–14.

Shoenberg, I. (1967). On spline functions. In O. Shisha (Ed.), *Proceedings of a symposium held at Wright-Patterson Air Force Base*, pp. 255–291. New York: Academic Press.

Turek, J. H. (1985). The northeast in a national context: background trends in population, income, and empoloyment. In H. W. Richardson and J. H. Turek (Eds.), *Economic prospects for the northeast*. Philadelphia: Temple University Press.

Van Nieuwerburgh, S. and P.-O. Weil (2010). Why has house price dispersion gone up? *Review of Economic Studies 77*, 1567–1606.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Appendix
(For online publication)

This appendix describes how we calculate gross migration flows using IRS data and our sources for the other data used in our analysis; the robustness of our empirical estimates and the evidence on multi-year migration discussed in the main text; one possible decentralization of the migration problem; why housing influences urban population dynamics; how we quantify the bias in using migration costs based on inter-state migration in a model of inter-city migration; and the methods used to solve the quantitative model.

A. Data

This section describes our data sources for our panel data on 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. As of that time, OMB defined 366 MSAs in the United States. Over our sample period, these MSAs account for about 80 percent of the aggregate population. We exclude all data from the New Orleans MSA from our study because of the disruption caused by Hurricane Katrina in 2005.

A.1. Calculating Gross Migration Flows with IRS Data

We construct data on gross MSA-level population inflows and outflows using county-to-county migration data based on tax records that is constructed by the Internal Revenue Service (IRS). These data are available annually from 1990 onwards on the IRS web site and are available from 1983 through through 1992 at the Inter-University Consortium for Political and Social Research (ICPSR) web site. The data 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. The beginning of the sample period for all our data is 1985, dictated by the availability of the IRS data. We do not use IRS data that is available prior to 1985 because counties are not identified by FIPS codes in these years.

For each year, the IRS reports the migration data using two files, one for outflows and one for inflows. These files cover the experience of 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.[32] We use the exemptions data.

[32] See `http://www.irs.gov/taxstats/article/0,,id=212683,00.html` for details.

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. We 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 365 MSAs, and for gross outflows the counties receiving the migrants are not restricted to be included in one of the 365 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, L_{it} as the number of people exiting, and $\bar{P}_t$ as all the people that did not move in or out. We compute the beginning of year population X_{it} and end of year population P_{it} as

$$\begin{aligned} X_{it} &= \bar{P}_{it} + L_{it} \\ P_{it} &= \bar{P}_{it} + A_{it} \end{aligned}$$

Net migration in MSA i during year t is therefore $P_{it} - X_{it} = A_{it} - L_{it}$. Note that due to births and deaths and foreign migration, P_{it} is typically less than the beginning of year population in the subsequent year, X_{it+1}. The corresponding gross arrival *rate*, gross departure *rate* and net migration *rate* are defined as

$$a_{it} = \frac{A_{it}}{X_{it}} \tag{A.1}$$

$$l_{it} \quad \frac{L_{it}}{X_{it}}$$

$$n_{it} = a_{it} - l_{it}. \tag{A.2}$$

A.2. Other data

Employment and wages are derived from Table CA4 of the "Regional Economic Accounts: Local Area Personal Income and Employment" as produced by the Bureau of Economic Analysis (BEA). 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. 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.[33] Since our detrending

[33] The BEA also reports a broader measure of employment called "total employment" (line 7010) but this is the sum of Wage and Salary Employment and Proprietors Employment (line

removes time effects, we do not adjust wages for overall inflation.

For house prices, we use the repeat-sales house price indexes produced by the Federal Home Finance Agency (FHFA) and available for download on that agency's web site. These data are quarterly; for each MSA, we construct an annual estimate as the average of all non-missing quarterly observations. For 11 larger 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. All MSAs have price-index data starting in 2001, but prior to that the time span of coverage varies by MSA. Table A.1 lists the number of MSAs with house price data we use in our study, by year

Table A.1: Number of MSAs for house price and residential investment data

Year	Number of MSAs	Year	Number of MSAs
1985	201	1997	364
1986	225	1998	364
1987	284	1999	364
1988	308	2000	364
1989	316	2001	365
1990	325	2002	365
1991	330	2003	365
1992	345	2004	365
1993	353	2005	365
1994	361	2006	365
1995	363	2007	365
1996	364		

The house price data are nominal indexes but this is unimportant as we log and first-difference the data, removing a scale factor that would covert the index to a level. What might be more important is the fact that the sample of MSAs changes over time and our detrending removes each year's cross-sectional average value (the "hat" from Section B.1 below.) Consider, as an example, a scenario where there are 5 MSAs used to compute an average for a few years, and then a 6th MSA is added to the sample. Consider only the original 5 MSAs. It is possible that the value of each of these MSAs relative to the average of the 5 original MSAs did not change; but, once the 6th MSA joins the sample, the value

7040). In CA4, Proprietors Income (line 70) is associated with Proprietors Employment. Since Proprietors Income includes some capital income, we use the narrower measures of employment and compensation in our study.

relative to the new average inclusive of the 6th MSA changes. We checked for this possibility by restricting cohorts of MSAs to be fixed over time at the entry year when computing sample averages. Referring to the previous example, for the original 5 MSAs, the sample average would always only be computed using those 5 MSAs; and, for the new 6th MSA, going forward the sample average would always only be computed using all 6 MSAs. This adjustment did not change any of our results for house prices or residential investment (see below), and so we chose the straightforward rule to always use all data.

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" available on the Department of Housing and Urban Development Web Site. We aggregate the county-level permits data to the MSA-level using the 2009 MSA definitions. We generate MSA-level indexes of nominal residential investment by multiplying the MSA-level house price indexes described earlier and building permits. This calculation yields the value of new housing. But, given we log and first difference the data, it also yields the relevant value for residential investment assume the share of new housing attributable to land is constant over time. Since we use house prices to construct residential investment, the number of observations by year is the same as for housing (see Table A.1) and all the caveats for the house price data apply here as well.

B. Detailed empirical analysis

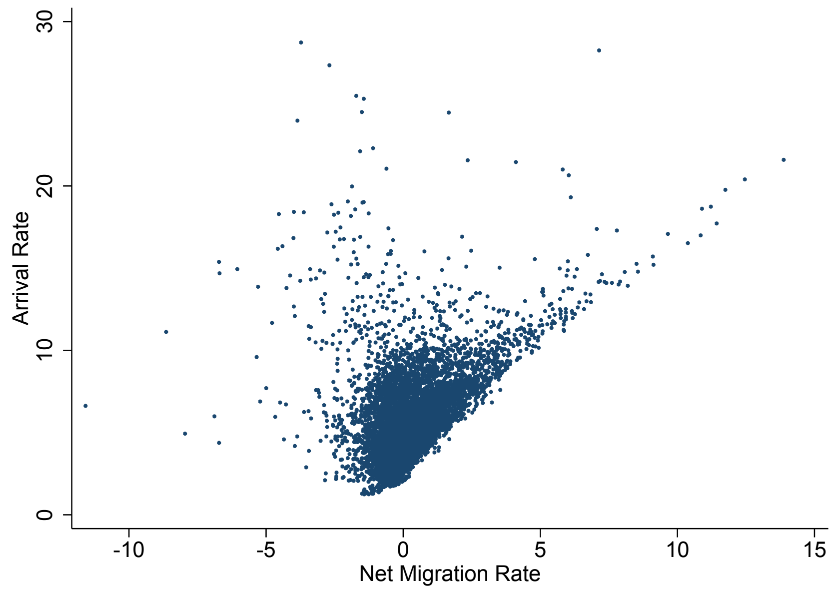
This appendix develops our findings about inter-city migration in more detail; shows how the choice of sample period affects our results; demonstrates our findings for multi-year migration; and shows that our key findings about inter-city migration extend to state-level data.

B.1. Effects of year and city variation on inter-city migration

Figure B.1 shows a scatter plot of raw arrival and net migration rates for all city-year observation in our baseline 1985-2007 sample. This plot is difficult to interpret because of the presence of a secular trend in gross migration and because there are systematic differences in gross migration rates depending on the size of the city (see Figure 1) and likely other factors, such as those considered by Coen-Pirani (2010), as well. It turns out that not taking into account these factors, particularly the city fixed effects, has a substantial impact on how the migration data is interpreted.

Figure B.2 shows the trends in gross migration. The observations for each year correspond to the cross-city average of the indicated gross migration rate,

Figure B.1: Raw city-level arrival versus net migration rates, 1985–2007



$\bar{z}_t$ for $z = a, l$, where

$$\bar{z}_t = \sum_{i=1}^{N} z_{it}$$

and N denotes the total number of MSAs in our data. In our data the average arrival rate always exceeds the corresponding departure rate because there is net in-migration to the cities in our sample from counties outside these cities and we include all US counties in our gross migration measures (see Appendix A.1).

Note that while the gross inter-*city* trends depicted in Figure B.2 are qualitatively consistent with the trends in gross inter-*state* migration studied by Kaplan and Schulhofer-Wohl (2015) the concepts underlying their measurement are a little different. Kaplan and Schulhofer-Wohl (2015) measure the likelihood that an individual drawn randomly from the entire population, regardless of location, will move (or have moved) to (from) another location (in their case state) in a given year. Our measure is the average across locations (in our case cities) of the likelihood that a person randomly drawn from a particular location will move (or have moved) to (from) another location. Population-weighted averages of location-specific gross rates would correspond to the concept studied by Kaplan and Schulhofer-Wohl (2015).

The trends evident in Figure B.2 suggest that time-series variation might

Figure B.2: Average city-level gross migration rates, 1985–2009

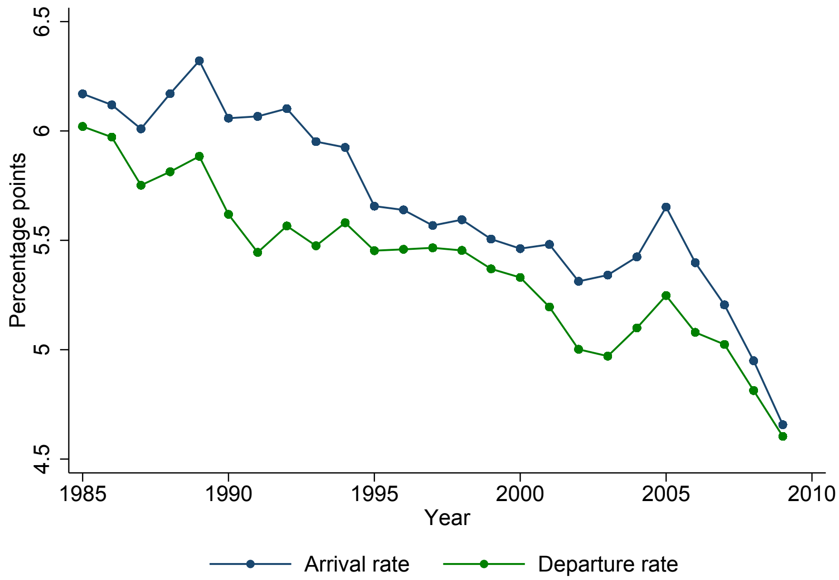
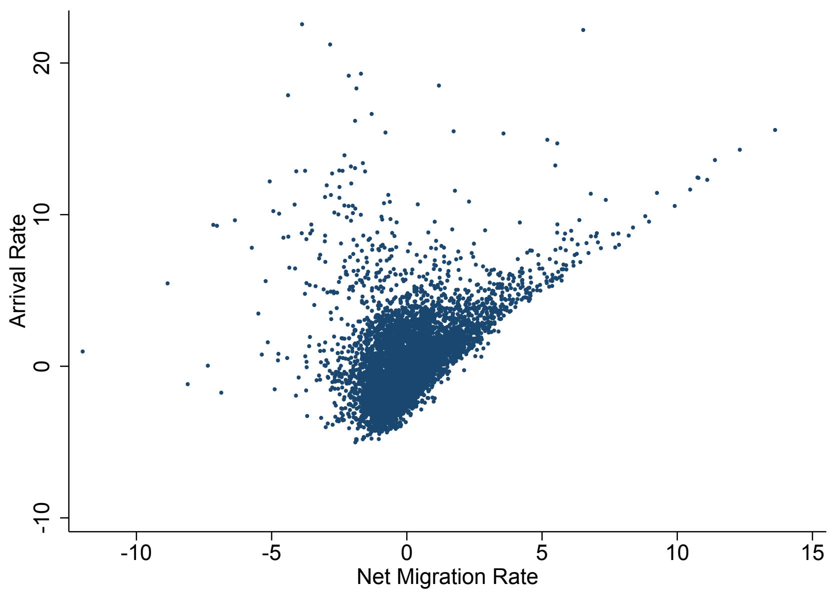


Figure B.3: Pooled arrival and net inter-city migration rates after removing year fixed effects only, 1985–2007



Note: Arrival rates correspond to $\tilde{a}_{it}$ and net migration rates correspond to $\tilde{n}_{it}$. See the text for the definition of these variables.

confound our interpretation of the data. We remove year effects as follows

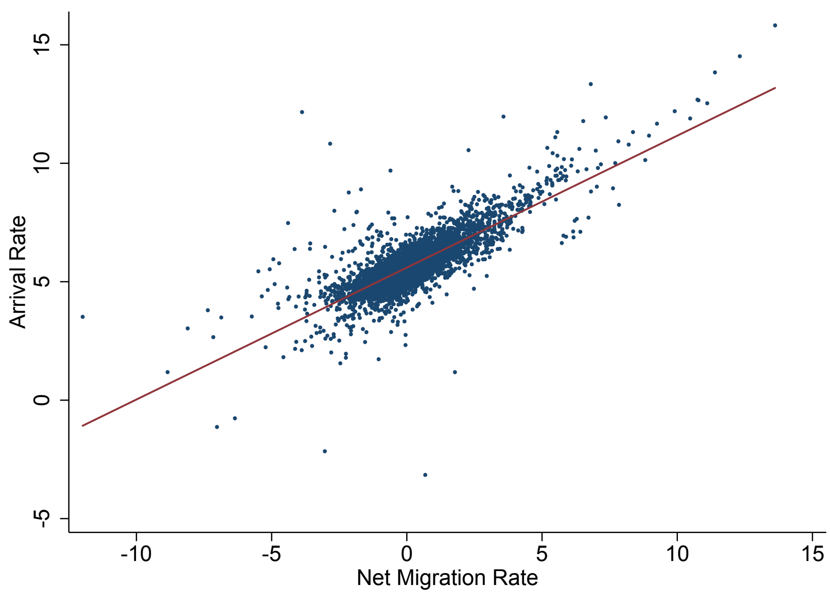
$$\tilde{a}_{it} = a_{it} - \bar{a}_t$$

and

$$\tilde{n}_{it} = \tilde{a}_{it} - \tilde{l}_{it}$$

where $\tilde{l}_{it}$ is defined analogously to $\tilde{a}_{it}$. Figure B.3 is a scatter plot of $\tilde{a}_{it}$ versus $\tilde{n}_{it}$. It appears that subtracting year effects alone has little impact on the relationship between arrival and net migration rates.

Figure B.4: Pooled arrival and net inter-city migration rates after removing year and city fixed effects, 1985–2007



Note: Arrival rates correspond to $\hat{a}_{it}$ and net migration rates correspond to $\hat{g}_{it}$. See the text for the definition of these variables.

Figure B.4 shows the importance of removing city fixed effects in gross migration. This figure depicts a scatter plot of $\hat{a}_{it}$ versus $\hat{n}_{it}$ where

$$\hat{a}_{it} = \tilde{a}_{it} - \frac{\sum_{\tau=1}^{T}\left(\tilde{a}_{i\tau} + \tilde{l}_{i\tau}\right)}{2T} + \frac{\sum_{\tau=1}^{T}\left(\bar{a}_{\tau} + \bar{l}_{\tau}\right)}{2T} \tag{B.1}$$

and

$$\hat{n}_{it} = \hat{a}_{it} - \hat{l}_{it}. \tag{B.2}$$

The departure rate $\hat{l}_{it}$ is defined analogously to $\hat{a}_{it}$. The specification of $\hat{a}_{it}$ subtracts the year fixed effect $\bar{a}_t$ and the city fixed effect $\sum_{\tau=1}^{T}\left(\tilde{a}_{i\tau}+\tilde{l}_{i\tau}\right)/(2T)$ from the raw arrival rate a_{it} and adds back the time-city mean of the average gross migration rates. Note that by specifying the city fixed effect as the mean of the time-series averages of the two gross migration rates (rather than the an arrival- or leaving-rate-specific fixed effect) ensures that we do not take out a city fixed effect in the net migration rate. Since Gibrat's law, which says population growth rates are independent city size, holds in our data there should not be a city fixed effect in net migration. Transforming the data in this way is also consistent with our model which satisfies Gibrat's law as well.

Figure B.4 reveals that, once city and year fixed effects are taken into account, there is a tight positive relationship between arrival and net migration rates. The red line shows the OLS fitted values. The individual observations line up remarkably well with the OLS regression line. Clearly not removing city fixed effects confounds within and cross-city variation. Cross-city gross migration rates tend to move together, due, at least in part, to the city-size effect. Figure B.4 strongly suggests that at the city level they tend to move in opposite directions because the slope of the OLS line is less than 1 (see below). Only by removing the city fixed effect does this pattern become apparent. Note that this pattern is evident when we consider the cities individually after removing year effects. In city-specific regressions of arrival rates on net migration rates, 86% of the coefficients lie between 0 and 1.

Table B.1 summarizes our findings thus far. It displays coefficients from regressions of arrival rates on net migration rates not taking into account any fixed effects (a_{it} versus n_{it}), accounting for year fixed effects only ($\tilde{a}_{it}$ versus $\tilde{n}_{it}$) and accounting for both year and city fixed effects ($\hat{a}_{it}$ versus $\hat{n}_{it}$). Both ordinary least squares (OLS) and least absolute deviations (LAD) or median regression estimates are shown. A coefficient greater than 1 would indicate that changes in net migration occur with changes in gross migration in the same direction, but with the arrival rate changing by more than the departure rate. Another possibility is that all changes in net migration occur along the arrival margin. This would be the implication of a coefficient insignificantly different from 1. A coefficient between 0 and 1 indicates that increases in net migration occur with arrival rates increasing and departure rates decreasing.

Consistent with the Figures B.1, B.3 and B.4, Table B.1 shows that considering the raw data or only taking into account year fixed effects leads to strikingly differently conclusions compared to if city fixed effects are accounted for. In the first two cases the OLS estimates are insignificantly different from 1. Mitigating the impact of outliers using LAD estimation changes this conclusion. The coefficients are significantly larger than 1 in these cases which indicates that both the arrival rate and departure rate margins are involved in variation in net migration

Table B.1: Impact of including year and city fixed effects, 1985–2007

Fixed effects included	Coefficient OLS	LAD
No fixed effects	.83***	1.19***
	(.12)	(.04)
Obs.	8395	8395
R^2	.18	.18
Year fixed effects only	.83***	1.18***
	(.12)	(.03)
Obs.	8395	8395
R^2	.18	.18
Year and city fixed effects	.56***	.56***
	(.007)	(.008)
Obs.	8395	8395
R^2	.66	.66

Note: Coefficients from regressions of arrival rates on net migration rates. Superscript *** indicates statistical significance at the 1 percent level. Standard errors are clustered by MSA. OLS denotes ordinary least squares and LAD denotes least absolute deviations (median regression).

but that these margins move in the same direction – net migration increases by both arrival and departure rates increasing, but arrival rates increasing by more. This is consistent with the positive correlation highlighted in Coen-Pirani (2010).

In the last panel we see that accounting for both year and city fixed effects changes our view about the nature of net migration. The coefficient is the same regardless of estimation method and in both cases equals 0.56 with a small standard error. These results suggest that increases in net migration within a city arise from increases in the arrival rate and decreases in the departure rate. Notice as well that the R^2 in these cases is much larger than if city fixed effects are not removed from the data – within cities there is a tight relationship between gross and net migration.

B.2. Alternative sample periods

We chose 2007 as the endpoint of our baseline sample because it is the last year before the collapse of the housing market. There is a substantial body of evidence, surveyed and built on in Modestino and Dennett (2013), that suggests a discrete change in gross migration occurred in the years immediately following the housing bust. The conventional explanation for this finding is that homeowners were unwilling to realize losses on their houses and so at the margin were less willing to move than would otherwise be the case. Since our model does not include this mechanism we think it is appropriate to exclude these data.

Still, we have data on migration running through 2009 so it is worth investigating the impact of adding the post-housing-bust years to our sample. While we are not aware of evidence suggesting the housing boom was related to unusual propensities to migrate it is possible that it was. This motivates considering the implications of ending our sample before the housing boom. Therefore now we consider the impact on our main findings of two alternatives to our baseline 1985–2007 sample: 1985–2004 and 1985–2009.

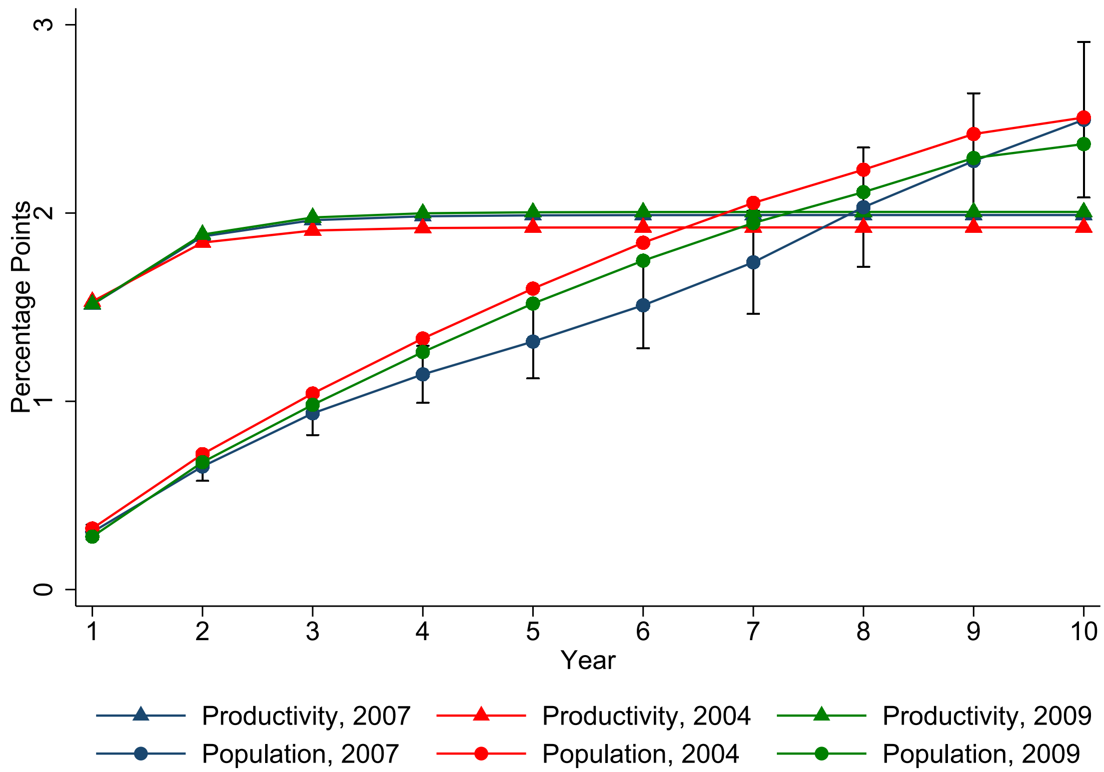
Table B.2: Inter-city migration estimates for different samples and estimation methods

	Coefficient	
Sample Period	OLS	LAD
1985–2007	.56***	.56***
	(.007)	(.008)
Obs.	8395	8395
R^2	.66	.66
1985–2004	.55***	.56***
	(.006)	(.007)
Obs.	7300	7300
R^2	.69	.69
1985–2009	.57***	.58***
	(.008)	(.008)
Obs.	9125	9125
R^2	.65	.65

Note: Superscript *** indicates statistical significance at the 1 percent level. Standard errors are clustered by MSA. OLS denotes ordinary least squares and LAD denotes least absolute deviations (median regression).

Table B.2 shows that estimates of the relationship between gross and net migration (after accounting for year and city fixed effects) are not materially impacted by the sample period chosen. The first panel reproduces the results from the bottom panel of Table B.1, our baseline estimates. The bottom two

Figure B.5: Responses of population and productivity based on different sample periods

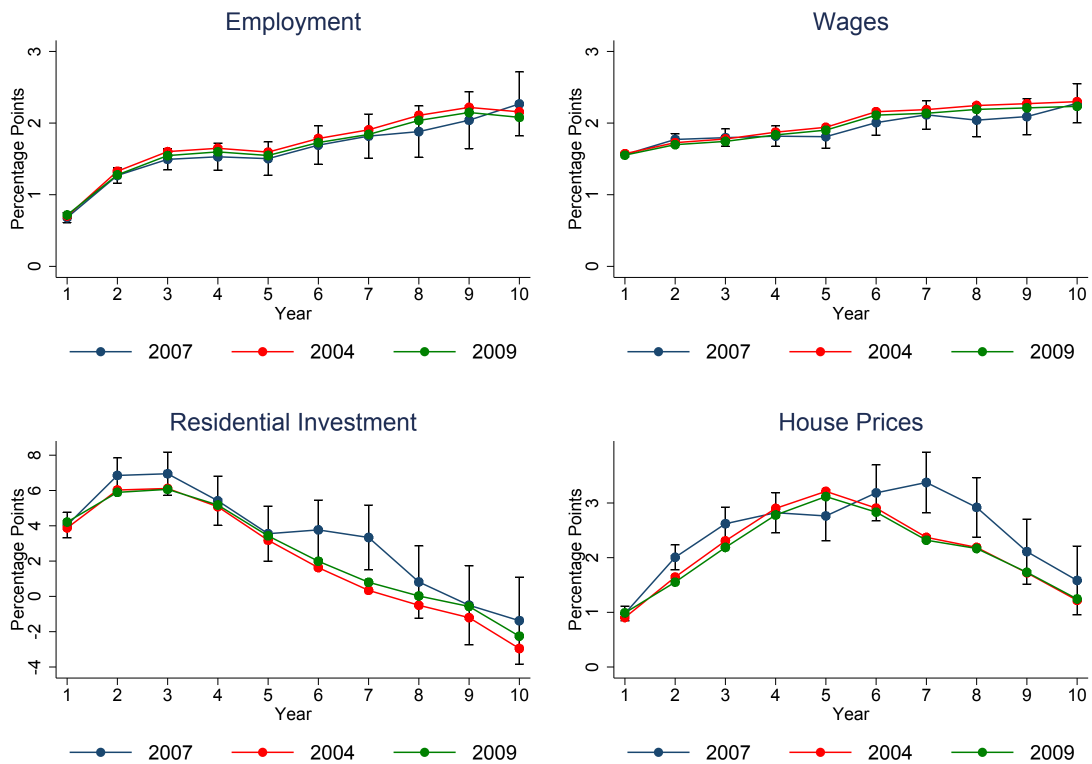


Note: Findings based on three sample period: 1985–2007, 1985–2004 and 1985–2009. Point estimates along with 2 standard error bands for population based on the 1985–2007 sample.

panels reveal that we get almost identical results with the shorter and longer sample periods.

Figures B.5, B.6 and B.7 display our baseline impulse response functions, considered in Sections 2.2 and 6.1, along with estimates corresponding to the two alternative samples. Except for the migration responses these plots indicate there is little effect of sample period on the estimates. Generally the responses for the shorter and longer samples fall within the 95% confidence bands associated with the baseline sample. When they do not fall within the confidence interval they do not miss by enough to overturn our conclusions that the model does a good job replicating the observations that were not directly targeted in the calibration. This also is true in the case of the migration responses for the shorter sample – excluding the housing boom as well as the housing bust does not change our conclusions.

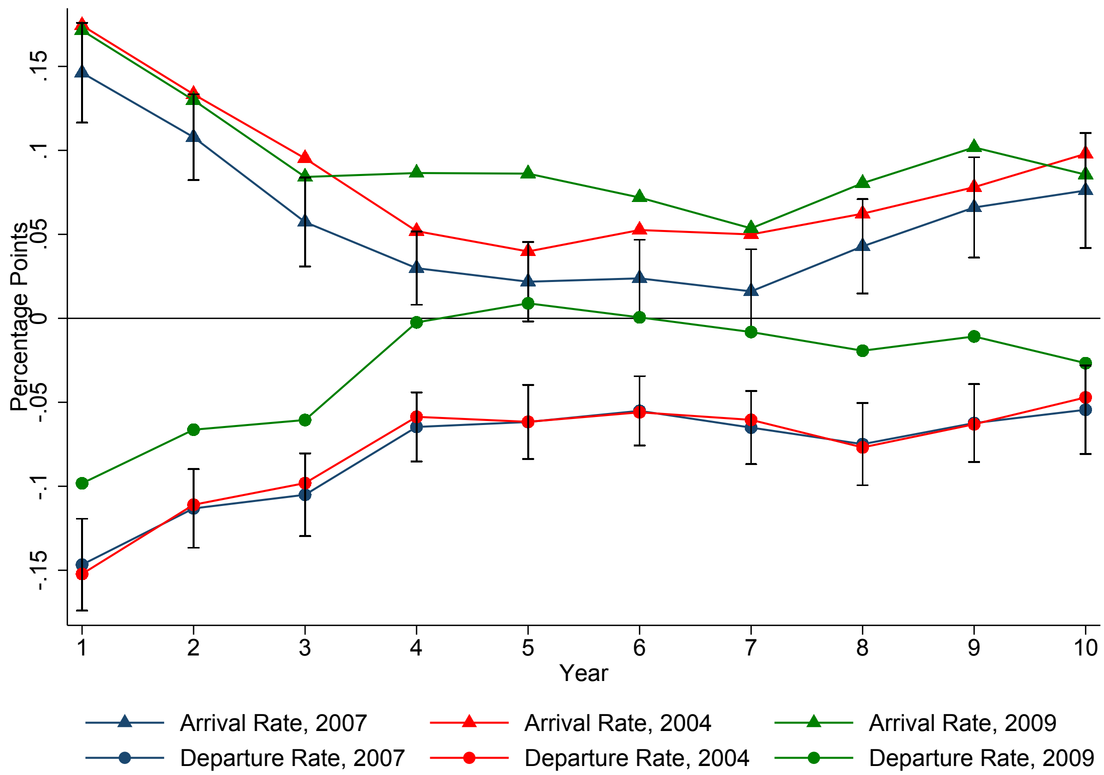
Figure B.6: Responses of labor and housing market variables based on different sample periods



Note: Findings based on three sample periods: 1985–2007, 1985–2004 and 1985–2009. Point estimates along with 2 standard error bands based on the 1985–2007 sample.

The longer sample migration responses, however, are significantly different from the baseline. Interestingly they are different in a way that confirms the empirical evidence reviewed and introduced in Modestino and Dennett (2013). To the extent that the housing bust deterred households from moving from their homes, its primary effect would be on the out-migration margin. Consistent with this hypothesis Figure B.7 shows that including the housing bust years in the sample period drives the departure rate response toward zero and the arrival rate away from zero. That is, during the housing bust net migration adjustments at the city-level occurred more on the arrival margin than the leaving margin compared to the pre-housing-bust period. Nevertheless the qualitative pattern is unaffected: it remains true regardless of sample period that net migration in response to productivity shocks adjusts by movements of arrival and departure rates in opposite directions.

Figure B.7: Responses of arrival and departure rates based on different sample periods



Note: Findings based on three sample periods: 1985–2007, 1985–2004 and 1985–2009. Point estimates along with 2 standard error bands based on the 1985–2007 sample.

B.3. Multi-year migration

In the main text we demonstrate that our model implies an affine relationship between a_{it+s} and n_{it+s}. Summarize this relationship, shown in equation (10), as

$$a_{it+s} \quad = \quad \alpha + \lambda n_{it+s}, \tag{B.3}$$

where λ and α should not be confused with parameters used in the main text (see equations A.1 and A.2 in Appendix A for the definitions of a_{it} and n_{it}). Denote by $g_{it,t+s}$ the gross growth rate of the population in MSA i from t to $t+s$:

$$g_{it,t+s} \quad = \quad \frac{X_{it+s}}{X_{it}} \tag{B.4}$$

and multiply each term in equation (B.3) by $g_{it,t+s}$:

$$a_{it+s}g_{it,t+s} \quad = \quad \alpha g_{it,t+s} + \lambda n_{it+s}g_{it,t+s}.$$

Since this equation holds for any s, it holds for the sum over s for $s = 0, 1, \ldots \tau - 1$. Therefore

$$\sum_{s=0}^{\tau-1} a_{it+s}g_{it,t+s} \quad = \quad \alpha \sum_{s=0}^{\tau-1} g_{it,t+s} + \lambda \sum_{s=0}^{\tau-1} n_{it+s}g_{it,t+s}. \tag{B.5}$$

The left-hand side of equation (B.5) has the interpretation of an arrival rate for a period of length τ years since

$$\sum_{s=0}^{\tau-1} a_{it+s}g_{it,t+s} \quad = \quad \frac{\sum_{s=0}^{\tau-1} A_{it+s}}{X_{it}}$$

Similarly, the term with λ in front equals the net-migration rate for a period of length τ years. Thus, our model implies a linear relationship between arrival rates and net migration rates for periods of any length in years with a slope coefficient that is identical to the slope coefficient for the one year horizon.

We can use this prediction to test the linearity hypothesis embedded in equation (B.3). To do this we estimate (B.5) for different values of τ and test the equality of the slope coefficients. To estimate equation (B.5), we first detrend a_{it+s}, g_{it+s} and n_{it+s}. We detrend a_{it+s} and n_{it+s} using equations (B.1) and (B.2). The detrended value of the gross population growth rate $g_{it,t+s}$, $\hat{g}_{it,t+s}$, is given by

$$\hat{g}_{it,t+s} = 1 + g_{it,t+s} - \sum_{i=1}^{N} g_{it,t+s}$$

That is the detrended gross population growth rate between years t and $t+s$ equals the gross population growth rate minus the average rate of gross population growth from t to $t+s$ across all MSAs plus 1. We use the BEA population data to calculate these population growth rates.

Define

$$\hat{a}_t^{\tau} \quad = \quad \sum_{s=0}^{\tau-1} \hat{a}_{it+s}\hat{g}_{it,t+s} - \bar{\alpha} \sum_{s=0}^{\tau-1} \hat{g}_{it,t+s}$$

$$\hat{n}_t^{\tau} \quad = \quad \sum_{s=0}^{\tau-1} \hat{n}_{it+s}\hat{g}_{it,t+s},$$

where $\bar{\alpha}$ equals the unconditional mean one-year gross migration rate from our data. We test the linearity hypothesis as follows. First we obtain the slope coefficients $\bar{\lambda}_\tau$ obtained from OLS regressions of $\hat{a}_t^\tau$ on $\hat{n}_t^\tau$ ($\bar{\lambda}_1 = \bar{\lambda} = .56$ is our estimate of the one year slope.) Second, we test the equality of the slope coefficients at different horizons, that is we test whether $\bar{\lambda}_\tau = \bar{\lambda}$ for different values of τ.

Our estimates of λ_τ for various values of τ are reported below in Table B.3, along with robust standard errors (clustered at the CBSA level) and R^2 values. The estimate of λ_τ hardly changes as we increase period length from one year ($\tau = 1$) to five years ($\tau = 5$). We cannot reject the hypothesis that the coefficients are identical at the 5% level.

Table B.3: Multi-year migration estimates for different horizons

τ	λ_τ	R^2
1	.56***	.68
	(.007)	
2	.56***	.70
	(.006)	
3	.56***	.72
	(.006)	
4	.55***	.74
	(.006)	
5	.55***	.75
	(.007)	

Note: Superscript *** indicates statistical significance at the 1 percent level. Standard errors are clustered by MSA.

It is instructive to present plots analogous Figure 2. To so we bin our data by deciles of the τ-year migration rates $\hat{n}_t^\tau$ and compute the mean net migration rate, $\hat{n}_t^\tau$, and the mean arrival rate, $\hat{a}_t^\tau$, for each bin. We also compute the mean τ-year departure rates where these departure rates are defined analogously to $\hat{a}_t^\tau$. Figures B.8 and B.9 plot the mean gross migration rates versus mean net migration rates by decile for $\tau = 3$ and $\tau = 5$. In these plots we have rescaled the gross migration rates by their corresponding unconditional means (17% for $\tau = 3$ and 28.5% for $\tau = 5$. Note that the slope coefficients are slightly different from the corresponding ones in Table B.3 since they are based on regressions using the binned data. These plots are strikingly similar to Figure 2.

Figure B.8: Gross migration rates by net Migration for $\tau = 3$

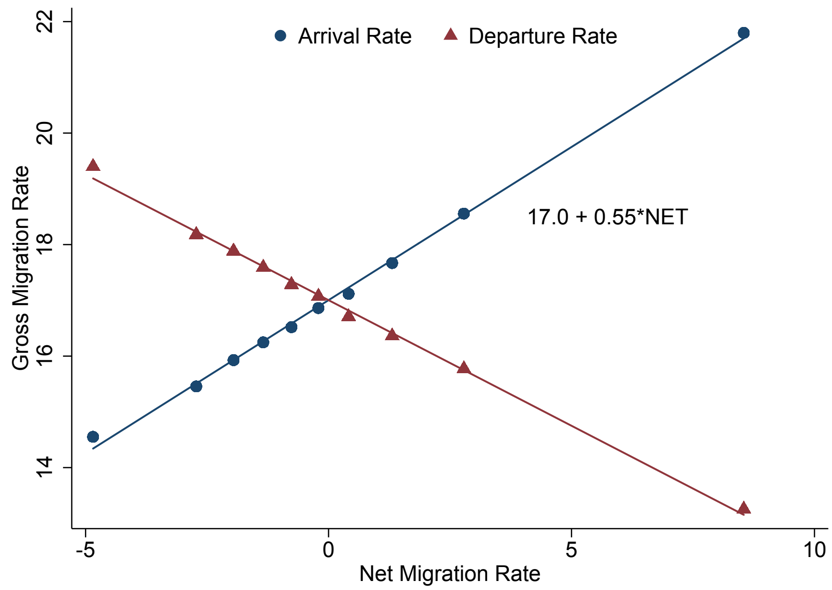
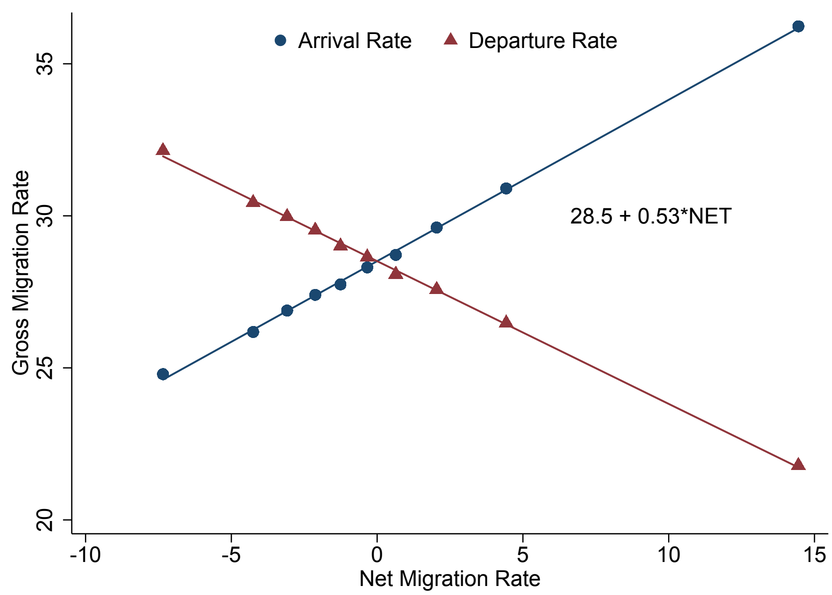


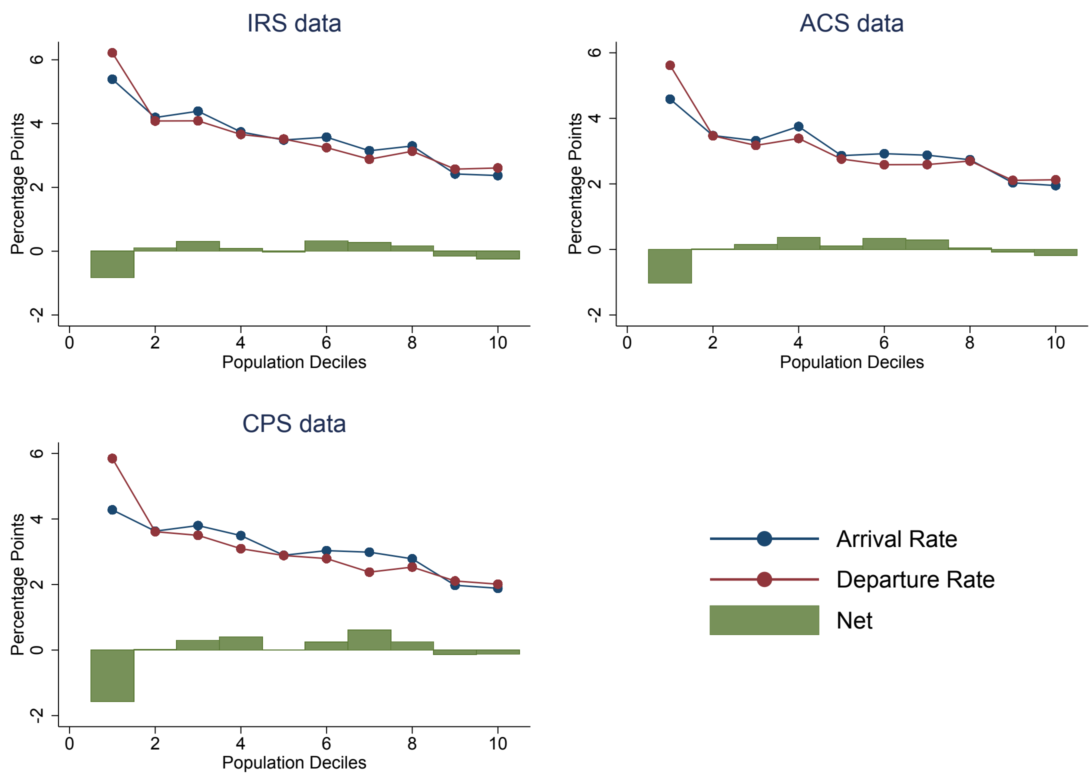
Figure B.9: Gross migration rates by net migration for $\tau = 5$



B.4. Gross and net inter-state migration

We now use three different data sources to demonstrate that our main findings for inter-city migration also hold for inter-state migration. Inter-state migration can be calculated using the IRS data by replacing cities with states in the methodology outlined in Appendix A.1. The March Supplement to the Current Population Survey (CPS) and the American Community Survey (ACS) also have information that can be used to calculate one-year gross and net inter-state migration. The samples in these latter two surveys are too small to study enough cities to provide a reliable view of inter-city migration.

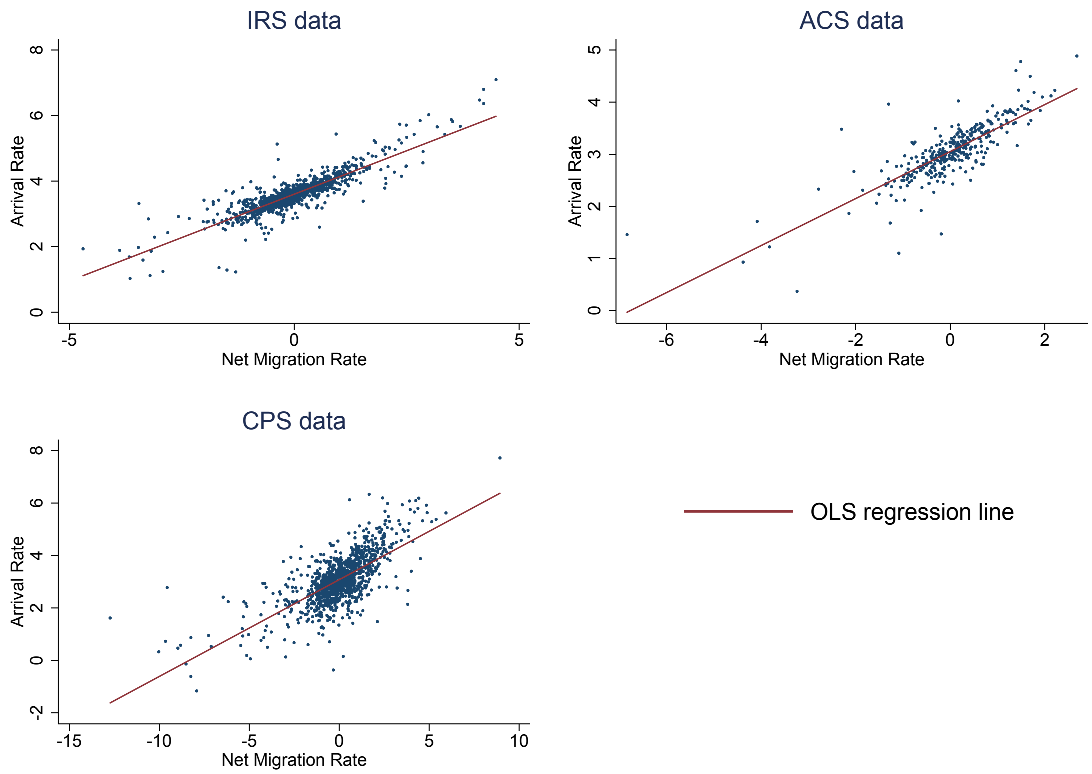
Figure B.10: Gross and net inter-state migration rates by population decile



In addition to verifying that the migration patterns we uncover are present in inter-state data as well, by comparing the IRS, CPS and ACS data on inter-state migration we can verify that the three data sources provide similar views of migration along the dimensions we are interested in. This is helpful because each data-set has their own idiosyncracies (see Kaplan and Schulhofer-Wohl (2015) and that paper's online appendix) and our results are more compelling if they are not just a feature of the IRS data we focus on.

We have CPS data covering the baseline sample period 1985-2007 (excluding 1995 because the CPS did not measure one-year migration that year) and ACS data for the period 2001-2007. In all cases we exclude Louisiana (due to Hurricane

Figure B.11: Gross and net inter-state migration rates for all year-state observations



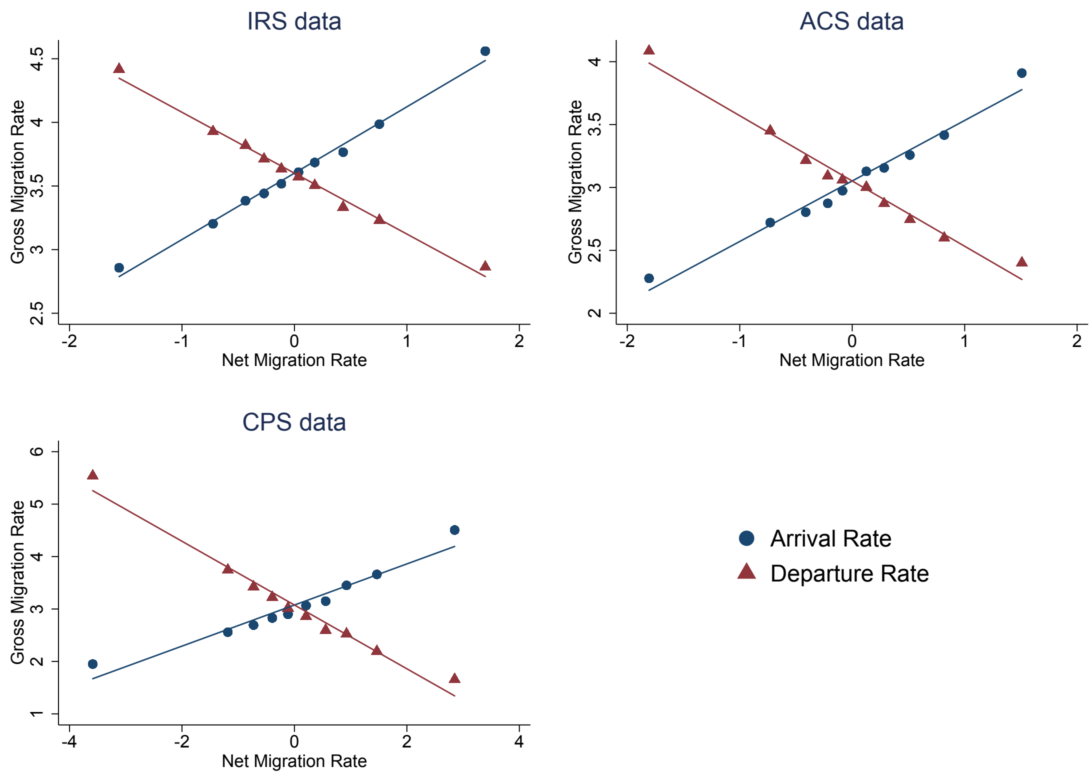
Katrina). Including the District of Columbia we have migration data for 50 "states" per year.

Figure B.10 displays the inter-state version of Figure 1 for each data-set. It demonstrates that gross inter-state migration is diminishing in the population of a state, just as gross inter-city migration is diminishing in the population of a city. This means that the relationship between inter-state gross and net migration will be confounded by cross-state variation.

Figure B.11 shows scatter plots of arrival and net migration states for all year-state observations in each data-set after removing year and state fixed effects. These data are defined analagously to $\hat{a}_{it}$ and $\hat{n}_{it}$ for cities. This figure demonstrates that all three data-sets display tight connections between arrival and net migration rates that are similar to the relationship we found for inter-city migration.

Figure B.12 plots the inter-state migration analogue of Figure 2 for each data-set. It shows that the striking linear relationship between gross and net migration that is present in inter-city migration is also present in inter-state migration.

Figure B.12: Gross versus net inter-state migration rates



Specifically, the arrival and net migration rates averaged by net migration decile line up almost exactly on the OLS regression lines in all three data-sets.

Table B.4 displays OLS and LAD estimates of linear regressions of inter-state arrival rates on net migration rates for the three data-sets over the longest sample period available for each data-set and for the longest sample period common to all three data-sets. This table shows that in all cases the coefficients lie between 0.34 and 0.53 and have confidence intervals far within the range $(0, 1)$. In other words the conclusion that increases in inter-city net migration occur via increases in arrival rates and decreases in departure rates extends to inter-state migration.

C. One Possible Decentralization of the Static Model

Decentralizing the static model's planning problem is useful for calibrating the quantitative model. The key challenge involved is how to treat guided trips. One valid approach is to have guided trips allocated entirely within the household through home production without any market interactions. We view guided trips in the model as an amalgam of both market and non-market activities and so we

Table B.4: Gross arrival versus net inter-state migration rates

	Longest Sample		2001–2007	
Data source	OLS	LAD	OLS	LAD
IRS	$.53^{***}$	$.53^{***}$	$.35^{***}$	$.35^{***}$
	(.01)	(.01)	(.05)	(.04)
Obs.	1150	1150	350	350
R^2	.77	.77	.47	.47
ACS	$.45^{***}$	$.49^{***}$	$.45^{***}$	$.49^{***}$
	(.03)	(.02)	(.03)	(.02)
Obs.	350	350	350	350
R^2	.66	.66	.66	.66
CPS	$.37^{***}$	$.41^{***}$	$.34^{***}$	$.34^{***}$
	(.03)	(.03)	(.03)	(.05)
Obs.	1100	1100	350	350
R^2	.49	.49	.36	.36

Note: Superscript *** indicates statistical significance at the 1 percent levels. Standard errors are clustered by state. OLS denotes ordinary least squares and LAD denotes least absolute deviations (median regression). The IRS and CPS longest samples are both 1985–2007; the longest sample for the ACS is 2001–2007.

think a more natural approach involves both activities. We now consider such a decentralization.

Markets are competitive. Firms in a city of type z hire labor at wage $w(z)$ and produce consumption goods to maximize profits. Household members initially located in a type-z city produce $a_m(z)$ guided trips to that city which they sell to prospective migrants at price $q(z)$. The household also home produces guided trips, denoted by $a_h(z)$. Let $m(z)$ denote the total number of guided trips to z-type cities purchased by household members in the market.

The representative household solves the following optimization problem

$$\max_{\substack{\{C,\Lambda,m(z),\\ a_m(z),a_h(z),\\ l(z),p(z)\}}} \quad \{\ln C - \int \left[\frac{A}{2} \left(\frac{a_m(z) + a_h(z)}{x} \right)^2 x$$

$$+\left(-\psi_1 \frac{l(z)}{x} + \frac{\psi_2}{2}\left(\frac{l(z)}{x}\right)^2\right) x \Bigg] d\mu - \tau\Lambda \Bigg\} \tag{C.1}$$

subject to:

$$C + \int q(z)\, m(z) d\mu = \int q(z)\, a_m(z) d\mu + \int w(z)\, p(z) d\mu + \int \Pi(z) d\mu \tag{C.2}$$

$$p(z) = x + m(z) + a_h(z) + \Lambda x - l(z),\ \forall z \tag{C.3}$$

$$\int l(z)\, d\mu = \int [m(z) + a_h(z) + \Lambda x]\, d\mu \tag{C.4}$$

along with non-negativity constraints on the choice variables. Equation (C.2) is the household's budget constraint where Π denotes profits from owning the firms. Equation (C.3) states that the population of a city after migration equals the initial population plus migrants from guided trips and undirected migration less the initial population that migrates out of the city. Finally, equation (C.4) states that the household members that migrate to cities must equal the number of household members that migrate out of cities.

The unique competitive equilibrium is defined in the usual way. Using the market clearing condition $m(z) = a_m(z)$ and the first order conditions of the household's problem we verify that a competitive equilibrium only determines the total number of guided trips into a city, $a_m(z) + a_h(z)$. The composition of these guided trips between market and non-market activities is left undetermined.

This particular decentralization makes it possible to calculate the total value of guided trips. In particular, as long as there are some guided trips purchased in the marketplace their total value is given by $q(z)a(z)$, with $a(z) = a_m(z) + a_h(z)$ and $q(z) = CAa(z)/x$.

D. Urban Population Dynamics with Housing

We expect housing to influence urban population dynamics for the reasons stated in the introduction: it is costly to build quickly, durable and immobile. Here we study a simple model to explain why these factors may be important. The model borrows the geography and production structure from Section 3. There are three differences with that model: individuals have a preference for housing services; to emphasize the role of housing, the model excludes migration frictions; and to study dynamics the model introduces infinitely lived households.

To analyze this simple model it is convenient to exploit the fact that the unique stationary competitive equilibrium can be obtained as the solution to a representative city planner's problem that maximizes local net surplus taking economy-wide variables as given, where the economy-wide variables are constrained to satisfy particular side conditions.[34] Since here we are only interested in the qualitative implications of housing we ignore the side conditions and study the city planner's problem assuming the economy-wide variables are exogenous.

Gross surplus in the representative city is given by

$$E_0 \sum_{t=0}^{\infty} \beta^t \left\{ s_t p_t^{\theta} + H \ln(\frac{h_t}{p_t}) p_t \right\}$$

where E_0 denotes the date $t = 0$ conditional expectations operator and $0 < \beta < 1$ is the household's time discount factor. Housing services are perfectly divisible so that each individual in the city enjoys h_t/p_t units of housing services where h_t denotes the local housing stock. The total utility individuals in the city derive from housing is given by $H \ln(h_t/p_t)p_t$, $H > 0$. The planner must give up $\eta > 0$ units of surplus for each individual it brings to the city and employs in the production of consumption goods.

Within this framework we consider the speed of adjustment of population to a one time permanent change to TFP. To be concrete, suppose the city is in a steady state at $t = 0$ with $s = s_0$ and then at date $t = 1$ it faces a one-time unanticipated permanent change in TFP to $s = s^*$. We consider the adjustment of population to this unanticipated change in TFP under three scenarios for housing.

In the first scenario the planner can rent housing services from other cities at the exogenous price r_h. This assumption is equivalent to assuming that housing is perfectly mobile across cities. Equilibrium in this scenario is characterized by the first order conditions for population and housing:

$$s_t \theta p_t^{\theta-1} + H \ln(\frac{h_t}{p_t}) = \eta + H; \tag{D.1}$$

$$H \frac{p_t}{h_t} = r_h. \tag{D.2}$$

Condition (D.1) shows that population adjusts to equate the marginal benefit of moving an additional individual to a city to the cost. The former is the sum of the individual's marginal product and housing services, while the latter includes the shadow cost of not moving the individual to another city and the housing services that are lost by the original inhabitants of the receiving city. The second condition equates the marginal utility of an extra unit of housing services with its cost. Replacing housing per individual in (D.1) using (D.2) yields

$$H \ln(\frac{H}{r_h}) - H + s_t \theta p_t^{\theta-1} = \eta. \tag{D.3}$$

[34]This is the same property we exploit when we analyze the quantitative model.

The key feature of (D.3) is that it does not include the housing stock, h_t. This means that after the unanticipated change in TFP population jumps immediately to its new permanent level $p = p^*$ found as the solution to (D.3) with $s_t = s^*$. So, when housing is perfectly mobile it is irrelevant for population dynamics.

Now assume the city is endowed with h_0 units of housing at $t = 0$ and that housing is immobile, meaning that it cannot be rented from or to any other city. In addition, suppose the change in TFP at $t = 1$ coincides with the onset of a potentially different exogenous path of the local housing stock satisfying

$$\ln h_t - \ln h^* = \rho_h^{t-1} (\ln h_0 - \ln h^*) \tag{D.4}$$

for $t \geq 1$, $0 \leq \rho_h < 1$, and h^* is the new long run level for h. Equilibrium population is determined by (D.1) conditional on (D.4). We consider two cases for this scenerio.

First suppose that the local, immobile housing stock does not change with TFP, that is $h^* = h_0$. From (D.1) after the change in TFP p jumps immediately to its new level given by the value p^* that solves this equation. The new long run level of population depends on h_0 but the speed of adjustment to p^* is the same as when housing is perfectly mobile. In other words the presence of local, immobile housing is not sufficient for housing to affect population dynamics.

The second case is the new long run level of housing changes with TFP, $h^* \neq h_0$. We approximate the transition of population to its new steady state in this case by log linearizing (D.1) around $\ln p^*$ and $\ln h^*$. This yields

$$\ln p_t - \ln p^* = \frac{H}{H + s^* \theta (1 - \theta) p^{*\theta - 1}} \rho_h^{t-1} (\ln h_0 - \ln h^*) .$$

In this case the speed of convergence of population to its new steady state p^* is directly related to the speed of convergence of housing to its new steady state through ρ_h. If the adjustment of housing is immediate, $\rho_h = 0$, then population's adjustment is instantaneous as in the case when $h^* = h_0$. If $0 < \rho_h < 1$ then population adjusts in proportion to the adjustment of housing.

We conclude that housing must be immobile and adjust slowly to changes in local productivity for it to affect population dynamics. It follows that a plausible quantitative analysis of urban population dynamics in response to TFP shocks must include endogenous immobile housing and include the possibility of its slow adjustment. Natural candidates for influencing the speed of adjustment of housing are construction depending on local resources and durability, which is the approach we take when formulating our quantitative model.

E. Inter-state and Inter-city Migration Costs in the Kennan and Walker (2011) Model

We now justify our conclusion that it is valid to apply Kennan and Walker (2011)'s estimate of migration costs in our environment. The argument is based on a calibrated model that incorporates the essence of the individual discrete choice problem studied by Kennan and Walker (2011) within an equilibrium setting.

There are N locations called cities. Each city i is associated with a wage that is fixed over time, w_i.[35] A person living in city i receives the wage and then receives a vector of i.i.d. preference shocks, one for each city including the person's current city, $e = (e_1, e_2, \ldots, e_N)$. After receiving the preference shocks, the person decides whether to move. The expected value of living in city i before the shocks are realized but after the wage is paid is

$$V_i \quad = \quad E\left[\max_{j\in\{1,\ldots,N\}}\left\{\frac{w_i}{\alpha} - \frac{c\left(i,j\right)}{\alpha} + e_j + \beta V_j\right\}\right]$$

Let s denote the state (a unique grouping of cities) containing city i and s' the state containing j. The moving cost function $c\left(i,j\right)$ is

$$c(i,j) = \begin{cases} 0, & \text{if } i = j \\ c_1, & \text{if } i \neq j \text{ and } s = s' \\ c_2, & \text{if } i \neq j \text{ and } s \neq s' \end{cases}$$

People pay no moving costs if they do not move, and in-state moving costs c_1 may be different than out-of-state moving costs c_2. Allowing c_1 to be different from c_2 is in the spirit of Kennan and Walker (2011)'s finding that moving costs increase with distance moved.[36]

Following Kennan and Walker (2011) we assume that the preference shocks are drawn from the Type 1 Extreme Value Distribution. Given a wage for each location w_i and the parameters of the model, α, β, c_1 and c_2, we compute the value functions using backwards recursion. We start with a guess of the expected value functions for every $j = 1, \ldots, N$. Call the current guess of the expected

[35] For simplicity we abstract from idiosyncratic wage shocks included by Kennan and Walker (2011). Kennan and Walker (2011) assume that individuals are finitely lived and only know the permanent component of wages of their current city and any city they have lived in previously. In an infinite horizon context individuals eventually live in every city and therefore have knowledge of the complete wage distribution.

[36] See the estimate of γ_1 in Table II on page 230 of their paper.

value function at location j as $\widehat{V}_j$. We then update the guess at each $i = 1, \ldots, N$

$$\widetilde{V}_i = \log \left\{ \sum_{j=1}^{N} \left[\exp \left(\frac{w_i}{\alpha} - \frac{c(i,j)}{\alpha} + \beta \widehat{V}_j \right) \right] \right\} + \zeta$$

where ζ is Euler's constant and $\widetilde{V}_i$ is the updated guess. We repeat this entire process until the expected value functions have converged, that is until $\widehat{V}_i$ is equal to $\widetilde{V}_i$ at each of the $i = 1, \ldots, N$ cities.

We set $N = 365$. For each city, we set w_i equal to the average wage in the corresponding MSA in 1990 (the year Kennan and Walker (2011) use to calculate average state wages) in thousands of dollars. For states that span multiple MSAs, we set the state s as the state where most of the population of the MSA lives in 1990.[37]

We assume that $\beta = 0.96$, leaving three parameters to be estimated: α, which scales the shocks into dollar equivalents, and the two moving costs c_1 and c_2. We estimate these parameters by targeting three moments: the average rate of individual migration across all MSAs, 4.47 percent, the average rate of across-state migration, 3.0 percent, and the average flow benefit scaled by average wage experienced by migrants, 1.9. For a worker moving from city i to city j, the flow benefit (scaled to dollars) is $\alpha (e_j - e_i) - c(i,j)$. Our target value of the average flow benefits of across-state movers scaled by average wage is taken from estimates produced by Kennan and Walker (2011) (see footnote 22 in the main text.) We use data from the IRS for 1990 to compute across-MSA and across-state migration rates. Our estimate of the across-state migration rate is almost identical to the estimate reported in Table VIII, page 239 by Kennan and Walker (2011) of 2.9 percent.

We compute all three moments analytically. The probability agents migrate to location j given their current location of i, $\gamma(j,i)$ has the straightforward expression

$$\gamma(j,i) = \frac{\exp\left(\frac{w_i}{\alpha} - \frac{c(i,j)}{\alpha} + \beta V_j\right)}{\sum_{k=1}^{N} \left[\exp\left(\frac{w_i}{\alpha} - \frac{c(i,k)}{\alpha} + \beta V_k\right)\right]}.$$

We construct the $N \times N$ matrix Γ, with individual elements $\gamma(j,i)$, and determine the steady state distribution of population across metro areas, the N-dimensional vector ρ, such that $\rho = \Gamma \rho$. Given ρ, we compute the probability of any move at

[37] Some MSAs span multiple states and this may introduce some error because within-MSA across-state moves that are truly within MSA will be misclassified as moves to a new labor market.

the steady-state population distribution as

$$\sum_{i=1}^{N} \rho(i) \left[\sum_{j \neq i} \gamma(j,i) \right]$$

and the probability of an across-state move as

$$\sum_{i=1}^{N} \rho(i) \left[\sum_{j \neq i, s' \neq s} \gamma(j,i) \right] .$$

For the third moment, it can be shown that the expected increase in continuation value from a worker choosing the optimal location as compared to an arbitrary location is a function of the probability the worker chooses the arbitrary location. For example, for a worker that optimally moves to location j, the expected increase in value, inclusive of flow utility and discounted future expected value, over remaining in the current location i is $-\log \gamma(i,i) / (1 - \gamma(i,i))$, see Kennan (2008). The expected increase in current flow payoff for all moves from j to i is therefore

$$\frac{-\log \gamma(i,i)}{1 - \gamma(i,i)} - \beta (V_j - V_i) .$$

The average of this second term across all moves, from i to all $j \neq i$, is

$$\frac{\sum_{j \neq i} \gamma(j,i) \beta (V_j - V_i)}{\sum_{k \neq i} \gamma(k,i)} = \frac{\sum_{j \neq i} \gamma(j,i) \beta (V_j - V_i)}{1 - \gamma(i,i)} .$$

The denominator is the probability a move occurs. Thus, conditional on moving, the average benefit of all moves that occur relative to staying put is

$$\sum_{i} \left(\frac{\rho(i)}{1 - \gamma(i,i)} \right) \left(-\log \gamma(i,i) - \sum_{j \neq i} \gamma(j,i) \beta (V_j - V_i) \right) .$$

We divide this expression by average wage (appropriately scaled), evaluated at the steady state: $\sum_{i} \rho(i) w_i / \alpha$.

We use the Nelder-Meade algorithm to search for parameters and we match our 3 target moments exactly. Our parameter estimates are $c_2 = 76.7$, $c_2 = 116.6$, and $\alpha = 17.6$. For reference, the mean wage at the steady state population distribution is 39.051 ($39,051). Our estimates of c_1 and c_2 imply that in-state and out-of-state moving costs are twice and three times average wages, respectively. These large costs generate low mobility rates in the face of large permanent wage differentials across metro areas. However, the estimated value of α implies that the mean and variance of the preference shocks are large. This large variance generates shocks large enough to induce people to move given the high costs of moving.

To determine the size of the bias in the Kennan and Walker (2011) estimates from using across-state moves, rather than across-MSA moves, we run 100 simulations of the model, simulating 600,000 people per MSA in each run. This generates approximately 27,000 moves to any MSA and 20,000 out of state moves for each MSA in the simulation. In each simulation run, we compute the economy-wide average flow benefits to across-state movers scaled by average wages. Averaged across the 100 simulations, the average benefit to across-state movers scaled by average wage is exactly 2.0, 0.1 higher than the average simulated benefits accruing to all movers. The bias is therefore 5%.[38] We find that a bias of this size does not affect our conclusions.

F. Solving the Quantitative Model

While representing the solution of the economy-wide social planner's as the solution to a city planner's problem plus side conditions is a huge simplification, computing the solution to the city planner's problem remains a nontrivial task.

The first difficulty is that the value function of the city planner's problem has two endogenous variables and two exogenous state variables where the latter two have very large domains. Each exogenous state variable takes values in a finite grid but this grid cannot be too coarse if the resulting discrete process is to represent the original AR(2) in a satisfactory way. To make the task of computing the value function manageable we used spline approximations.

Cubic spline interpolation is usually used in these cases. A difficulty with these methods is that they do not necessarily preserve the shape of the original function, or if they do (as with Schumacher shape-preserving interpolation) it is somewhat difficult to compute. For these reasons, we use a local method that does not interpolate the original function but that approximates it while preserving shape (monotonicity and concavity). An additional benefit is that it is extremely simple to compute (there is no need to solve a system of equations). The method is known as the Shoenberg's variation diminishing spline approximation. It was first introduced by Shoenberg (1967) and is described in a variety of sources (e.g. Lyche and Morken (2011)).

[38]Measured across the 100 runs, the standard deviation of the percent of the bias is 0.2%, the minimum bias is 4.6% and the maximum bias is 5.7%.

For a given continuous function f on an interval $[a, b]$, let p be a given positive integer, and let $\tau = (\tau_1, ..., \tau_{n+p+1})$ be a knot vector with $n \geq p+1$, $a \leq \tau_i \leq b$, $\tau_i \leq \tau_{i+1}$, $\tau_{p+1} = a$ and $\tau_{n+1} = b$. The *variation diminishing spline approximation of degree p to f* is then defined as

$$S_p(x) = \sum_{j=1}^{n} f\left(\tau_j^*\right) B_{jp}(x)$$

where $\tau_j^* = (\tau_{j+1} + ... + \tau_{j+p})/p$ and $B_{jp}(x)$ is the jth B-spline of degree p evaluated at x. The B-splines are defined recursively as follows

$$B_{jp}(x) = \frac{x - \tau_j}{\tau_{j+p} - \tau_j} B_{j,p-1}(x) + \frac{\tau_{j+1+p} - x}{\tau_{j+1+p} - \tau_{j+1}} B_{j+1,p-1}(x)$$

with

$$B_{j0}(x) = \begin{cases} 1, \text{ if } \tau_j \leq x < \tau_{j+1} \\ 0, \text{ otherwise} \end{cases}$$

As already mentioned, this spline approximation preserves monotonicity and concavity of the original function f (e.g. Lyche and Morken (2011), Section 5.2). The definition of variation diminishing splines is easily generalized to functions of more than one variable using tensor products (e.g. Lyche and Morken (2011), Section 7.2.1). These properties greatly simplify the value function iterations of the city planner's problem and they should prove useful in a variety of other settings. In our actual computations we worked with an approximation of degree $p = 3$.

An additional complication involves the return function of the city planner's problem. Conditional on the current states (h, x, s, s_{-1}) and future states (h', x'), evaluating the one period return function of the city planner requires solving a nonlinear system of equations in $(n_y, n_h, k_y, k_h, b_r, b_h)$ allowing for the possibility that the constraint $n_y + n_h \leq x'$ may bind. This is not a hard task. However, doing this for every combination of (h, x, s, s_{-1}) and (h', x') considered in solving the maximization problem at each value function iteration would slow down computations quite considerably. For this reason, we chose to construct a cubic variation-diminishing spline approximation to the return function $R(h, x, s, , s_{-1}, h', x')$ once, before starting the value function iterations, and use this approximation instead. In practice, for each value of (h, x, s, s_{-1}) we used a different knot vector for h' and x' to gain accuracy of the return function over the relevant range.

Performing the maximization over (h', x') for each value of (h, x, s, s_{-1}) at each value function iteration is a well behaved problem given the concavity of the spline approximations to the return function and the next period value function.

There are different ways of climbing such a nice hill in an efficient way. In our case, given that we could offload computations into two Tesla C2075 graphic cards (with a total of 896 cores), we used the massively parallel capabilities of the system to implement a very simple generalized bisection method. Essentially for each value of (h, x, s, s_{-1}) we used a block of 16×16 threads to simultaneously evaluate 16×16 combinations of (h', x') over a predefined square. We then zoom to the smallest square area surrounding the highest value and repeat. In practice, a maximum would be found after only three or four passes.

Statistics under the invariant distribution were computed using Monte Carlo simulations. This part of the computations was also offloaded to the graphic cards to exploit their massively parallel capabilities. To avoid costly computations similar to those encountered in the evaluation of the return function, cubic spline approximations were used for all decision rules.

Speeding up the solution to the city planner's problem and Monte Carlo simulations was crucial since finding solutions (Y, C, Λ, η) to the side conditions requires solving the city planner's problem and simulating its solution several times.

The source code, which is written in CUDA Fortran, is available upon request. Compiling it requires the PGI Fortran compiler. Running it requires at least one NVIDIA graphic card with compute capability higher than 2.0.