Econometrica Supplementary Material

SUPPLEMENT TO "MACROECONOMIC IMPLICATIONS OF AGGLOMERATION"

(*Econometrica*, Vol. 82, No. 2, March 2014, 731–764)

BY MORRIS A. DAVIS, JONAS D. M. FISHER, AND TONI M. WHITED[1]

This document provides supplementary background material for the main paper. It discusses (i) the growth model, (ii) the data, (iii) deriving the moment conditions that underlie estimation, (iv) measuring the impact of agglomeration on per capita consumption growth, (v) solving the model, (vi) standard errors of our estimates, (vii) Monte Carlo analysis of our estimation strategy, (viii) proofs of equilibrium existence and uniqueness for versions of the model without housing, (ix) several perturbations to our estimation, and (x) how we verify numerically that the Luttmer (2007) property holds.

APPENDIX A: FIRST ORDER CONDITIONS, PRICES, AND GROWTH

THIS SECTION DESCRIBES DETAILS about the growth model that were omitted from the main text. These include (i) the first order conditions of the planning problem, (ii) how to map the Lagrange multipliers for this problem into competitive equilibrium prices, (iii) how the planner's first order conditions relate to those of the agents' in the competitive equilibrium, (iv) the stationarity inducing transformation of the growing economy, and (v) how to calculate the increase in steady state consumption and housing required to compensate households for not having the growth due to local agglomeration.

A.1. *The Model Without Growth*

The competitive equilibrium can be found as the solution to an optimization problem with side conditions. Idiosyncratic technology z_t evolves as a stationary discrete Markov chain. Let $q_t(z^t)$ denote the time t distribution of cities across productivity histories z^t and let $Q(z, z')$ denote the probability that $z_{t+1} = z'$ conditional on $z_t = z$. The planner's problem is given by

$$\max_{\substack{\{C_t, K_{bt+1}, K_{st+1}, y(z^t), l_b(z^t), l_h(z^t),\\ k_b(z^t), k_s(z^t), k_{ft+1}(z^t), n(z^t), h(z^t)\}_{t=0}^{\infty}}} \left[\sum_{t=0}^{\infty} \beta^t \ln C_t \right.$$
$$\left. + \psi \sum_{t=0}^{\infty} \beta^t \sum q_t(z^t) n(z^t) \ln \frac{h(z^t)}{n(z^t)} \right]$$

[1]We are grateful to Bill Johnson and Randal Verbrugge for helping us to understand the evolution of the MSA definitions used at the Bureau of Labor Statistics. Any views expressed herein are those of the authors and do not necessarily represent those of the Federal Reserve Bank of Chicago or the Federal Reserve System.

 DOI: 10.3982/ECTA9029

subject to

$$C_t + P_{bt}\left[K_{bt+1} - (1-\kappa_b)K_{bt}\right] + P_{st}\left[K_{st+1} - (1-\kappa_s)K_{st}\right] + P_{ft}\sum q_t(z^t)\left[k_{ft+1}(z^t) - (1-\kappa_f)k_{ft}(z^{t-1})\right] \leq \left[\sum q_t(z^t)y(z^t)^{\eta}\right]^{1/\eta}, \tag{S1}$$

$$y(z^t) \leq A_t^{(1-\alpha)\phi} z_t^{(1-\alpha)\phi}\left[x(z^t)\right]^{(\lambda-1)/\lambda} l_b(z^t)^{1-\phi} k_b(z^t)^{\alpha\phi} n(z^t)^{(1-\alpha)\phi}, \quad \forall z^t, \tag{S2}$$

$$h(z^t) \leq l_h(z^t)^{1-\omega} k_s(z^t)^{\omega}, \quad \forall z^t, \tag{S3}$$

$$l_h(z^t) + l_b(z^t) \leq k_{ft}(z^{t-1})^{\zeta}, \quad \forall z^t, \tag{S4}$$

$$\sum q_t(z^t)k_b(z^t) \leq K_{bt}, \tag{S5}$$

$$\sum q_t(z^t)k_s(z^t) \leq K_{st}, \tag{S6}$$

$$\sum q_t(z^t)n(z^t) \leq 1, \tag{S7}$$

and $K_{b0}, K_{s0}, k_f(z_0), \{A_t, P_{bt}, P_{st}, P_{ft}, z^t\}_{t=0}^{\infty}$, and $x(z^t)$ given. Competitive equilibrium allocations are obtained as a solution to this optimization problem such that $x(z^t) = y(z^t)/l_b(z^t)$. Prices that correspond to an equilibrium are easy to obtain from the constraint's Lagrange multipliers. We now show how to do this and how to relate the planners first order conditions to those of individual agents in the competitive equilibrium.

Let the Lagrange multipliers for the above constraints be $\beta^t\pi_t$, $\beta^t\pi_t q_t(z^t) \times p_y(z^t)$, $\beta^t\pi_t q_t(z^t) r_h(z^t)$, $\beta^t\pi_t q_t(z^t) r_l(z^t)$, $\beta^t\pi_t r_{bt}$, $\beta^t\pi_t r_{st}$, and $\beta^t\pi_t\theta_t$. Then the first order conditions for C_t, K_{bt+1}, K_{st+1}, $k_{ft+1}(z^t)$, $y(z^t)$, $l_b(z^t)$, $k_b(z^t)$, $n(z^t)$, $h(z^t)$, $l_h(z^t)$, and $k_s(z^t)$ are

$$\frac{1}{C_t} = \pi_t, \tag{S8}$$

$$\pi_t P_{bt} = \beta\pi_{t+1}\left[r_{bt+1} + P_{bt+1}(1-\kappa_b)\right], \tag{S9}$$

$$\pi_t P_{st} = \beta\pi_{t+1}\left[r_{st+1} + P_{st+1}(1-\kappa_s)\right], \tag{S10}$$

$$\pi_t P_{ft} = \beta\pi_{t+1}\sum_{z_{t+1}}\left[r_{ft+1}(z^{t+1}) + P_{ft+1}(1-\kappa_f)\right]Q(z_t, z_{t+1}), \tag{S11}$$

$$p_y(z^t) = Y_t^{1-\eta} y(z^t)^{\eta-1}, \tag{S12}$$

(S13) $$r_l(z^t) = p_y(z^t)(1-\phi)z_t^{(1-\alpha)\phi}[x(z^t)]^{(\lambda-1)/\lambda} l_b(z^t)^{-\phi} k_b(z^t)^{\alpha\phi} n(z^t)^{(1-\alpha)\phi},$$

(S14) $$r_b(z^t) = p_y(z^t)\alpha\phi z_t^{(1-\alpha)\phi}[x(z^t)]^{(\lambda-1)/\lambda} \times l_b(z^t)^{1-\phi} k_b(z^t)^{\alpha\phi-1} n(z^t)^{(1-\alpha)\phi},$$

(S15) $$\pi_t\theta_t = \psi \ln \frac{h(z^t)}{n(z^t)} - \psi + \pi_t w(z^t),$$

(S16) $$\frac{\psi}{h(z^t)} = \pi_t r_h(z^t),$$

(S17) $$r_l(z^t) = (1-\omega) r_h(z^t) l_h(z^t)^{-\omega} k_h(z^t)^{\omega},$$

(S18) $$r_s(z^t) = \omega r_h(z^t) l_h(z^t)^{1-\omega} k_s(z^t)^{\omega-1},$$

where

(S19) $$w(z^t) = p_y(z^t)(1-\alpha)\phi z_t^{(1-\alpha)\phi}[x(z^t)]^{(\lambda-1)/\lambda} \times l_b(z^t)^{1-\phi} k_b(z^t)^{\alpha\phi} n(z^t)^{(1-\alpha)\phi-1},$$

(S20) $$Y_t = \left[\sum q(z^t) y(z^t)^{\eta}\right]^{1/\eta},$$

(S21) $$Q(z_t, z_{t+1}) = \frac{q_{t+1}(z^t, z_{t+1})}{q_t(z^t)},$$

(S22) $$r_f(z^t) = r_l(z^t)\zeta k_{ft}(z^{t-1})^{\zeta-1}.$$

In the competitive equilibrium $\{P_{xt}, x = b, s, f\}$, p_y, $\{r_x, x = b, s, f, l, h\}$, and w correspond to investment prices, intermediate good prices, rental rates, and wages. Under this interpretation, we can relate the planner's first order conditions to those of the individual agents described in Section 2.2 and referred to in Sections 3.1 and 3.2. Equations (S9)–(S11), after substituting for π using (S8), correspond to the representative household's first order conditions for capital accumulation. Equation (S12) corresponds to the first order condition of the final good producer for intermediate input demand. Equations (S13), (S14), and (S19) correspond to the intermediate good producers' first order conditions for finished land, business capital, and labor. Equation (S16), after substituting for π using (S8), corresponds to the household's first order condition for housing in each location. Equations (S17) and (S18) correspond to the first order conditions of housing service providers for finished land use and rental of residential structures. Equation (S22) corresponds to the first order condition of finished land service providers (landlords) for infrastructure use.

A.2. *The Model With Growth*

With growth, that is under the assumptions

$$
\begin{aligned}
&P_{xt} = \gamma_x^{-t}, \quad x = s, b, f, \\
&A_t = \gamma_a^t, \\
&N_t = \gamma_n^t,
\end{aligned}
$$

we need to obtain the mapping from the growing economy to a stationary planning problem. The mapping is driven by the balanced growth expressions derived in the main text.

The model's quantities are transformed as

$$
\begin{aligned}
&g_c^t \bar{C}_t = \frac{C_t}{N_t}, \quad g_x^t \bar{K}_{xt} = \frac{K_{xt}}{N_t}, \quad x = b, s, \quad g_c^t \frac{\bar{y}(z^t)}{\bar{n}(z^t)} = \frac{y(z^t)}{n(z^t)}, \\
&g_l^t \frac{\bar{l}_b(z^t)}{\bar{n}(z^t)} = \frac{l_b(z^t)}{n(z^t)}, \quad g_x^t \frac{\bar{k}_x(z^t)}{\bar{n}(z^t)} = \frac{k_x(z^t)}{n(z^t)}, \quad x = b, s, \\
&\gamma_x^{-t} \bar{P}_{xt} = P_{xt}, \quad x = b, s, f, \quad g_l^t \frac{\bar{l}_h(z^t)}{\bar{n}(z^t)} = \frac{l_h(z^t)}{n(z^t)}, \\
&g_f^t \frac{\bar{k}_f(z^{t-1})}{\bar{n}(z^t)} = \frac{k_f(z^{t-1})}{n(z^t)}, \quad \bar{n}(z^t) = \frac{n(z^t)}{N_t}, \quad \frac{g_c^t}{g_l^t} \bar{x}(z^t) = x(z^t), \\
&g_h^t \frac{\bar{h}(z^t)}{\bar{n}(z^t)} = \frac{h(z^t)}{n(z^t)}.
\end{aligned}
$$

All of the growth rates in these expressions are derived in the main text except for g_h. This growth rate equals $[\gamma_n^{\zeta-1} g_f^{\zeta}]^{1-\omega} g_s^{\omega}$.

The multipliers and prices are transformed as

$$
\begin{aligned}
&\bar{r}_{xt} = \gamma_x^t r_{xt}, \quad x = b, s, f, \\
&\bar{r}_{lt} = g_{p_l}^{-t} r_{lt}, \\
&\bar{r}_{ht} = g_{p_h}^{-t} r_{hlt}, \\
&\bar{\theta}_t = g_c^{-t} \theta_t, \\
&\bar{\pi}_t = g_c^t \gamma_n^t \pi_t, \\
&\bar{w}_t(z^t) = g_c^{-t} w(z^t).
\end{aligned}
$$

Replacing the growing variables in the original planning problem using these transformations, one can show the planning problem reduces to

$$\max_{\substack{\{\bar{C}_t,\bar{K}_{bt+1},\bar{K}_{st+1},\\ \bar{y}(z^t),\bar{l}_b(z^t),\bar{l}_h(z^t),\\ \bar{n}(z^t),\bar{k}_b(z^t),\bar{k}_s(z^t),\\ \bar{k}_f(z^t),\bar{h}(z^t)\}_{t=0}^{\infty}}} \left[\sum_{t=0}^{\infty}\beta^t \ln \bar{C}_t + \sum_{t=0}^{\infty}\beta^t \sum q_t(z^t)\bar{n}(z^t)\psi \ln \frac{\bar{h}(z^t)}{\bar{n}(z^t)} + \sum_{t=0}^{\infty}\beta^t \sum q_t(z^t)\bar{n}(z^t)\psi \ln g_h^t\right]$$

subject to

$$\begin{aligned}
&\bar{C}_t + \bar{P}_{bt}\big[\gamma_n g_b \bar{K}_{bt+1} - (1-\kappa_b)\bar{K}_{bt}\big] \\
&\quad + \bar{P}_{st}\big[\gamma_n g_s \bar{K}_{st+1} - (1-\kappa_b)\bar{K}_{st}\big] \\
&\quad + \bar{P}_{ft}\sum q_t(z^t)\big[\gamma_n g_f \bar{k}_{ft+1}(z^t) - (1-\kappa_f)\bar{k}_{ft}(z^t)\big] \\
&\leq \Big[\sum q_t(z^t)\big(\bar{y}(z^t)\big)^{\eta}\Big]^{1/\eta}, \\
&\bar{y}(z^t) \leq z_t^{(1-\alpha)\phi}\big[\bar{x}(z^t)\big]^{(\lambda-1)/\lambda}\big[\bar{l}_b(z^t)\big]^{1-\phi}\big[\bar{k}_b(z^t)\big]^{\alpha\phi}\big[\bar{n}(z^t)\big]^{(1-\alpha)\phi}, \quad \forall z^t, \\
&\bar{h}(z^t) \leq \big[\bar{l}_h(z^t)\big]^{1-\omega}\big[\bar{k}_s(z^t)\big]^{\omega}, \quad \forall z^t, \\
&\bar{l}_h(z^t) + \bar{l}_b(z^t) \leq \big[\bar{k}_{ft}(z^t)\big]^{\zeta}\gamma_n^{1-\zeta}, \quad \forall z^t, \\
&\sum q_t(z^t)\bar{k}_b(z^t) \leq \bar{K}_{bt}, \\
&\sum q_t(z^t)\bar{k}_s(z^t) \leq \bar{K}_{st}, \\
&\sum q_t(z^t)\bar{n}(z^t) \leq 1,
\end{aligned}$$

and $\bar{K}_{b0}$, $\bar{K}_{s0}$, $\bar{k}_f(z_0)$, $\{\bar{P}_{bt}, \bar{P}_{st}, \bar{P}_{ft}, z^t\}_{t=0}^{\infty}$, and $\bar{x}(z^t)$ given. Competitive equilibrium allocations for the growing economy are obtained in two steps. First we find the solution to the transformed problem such that $\bar{x}(z^t) = \bar{y}(z^t)/\bar{l}_b(z^t)$. The second step translates the stationary allocations and prices to their growing counterparts using the transformations described above.

Let the Lagrange multipliers for the above constraints be $\beta^t\bar{\pi}_t$, $\beta^t\bar{\pi}_t q_t(z^t) \times p_y(z^t)$, $\beta^t\pi_t q_t(z^t)\bar{r}_h(z^t)$, $\beta^t\pi_t q_t(z^t)\bar{r}_l(z^t)$, $\beta^t\pi_t\bar{r}_{bt}$, $\beta^t\pi_t\bar{r}_{st}$, and $\beta^t\pi_t\bar{\theta}_t$. Then the first order conditions for $\bar{C}_t$, $\bar{K}_{bt+1}$, $\bar{K}_{st+1}$, $\bar{k}_{ft+1}(z^t)$, $\bar{y}(z^t)$, $\bar{l}_b(z^t)$, $\bar{k}_{bt}(z^t)$, $\bar{n}(z^t)$,

$\bar{h}(z^t)$, $\bar{l}_h(z^t)$, and $\bar{k}_s(z^t)$ are

$$\frac{1}{\bar{C}_t} = \bar{\pi}_t,$$

$$\gamma_n g_b \bar{\pi}_t \bar{P}_{bt} = \beta \bar{\pi}_{t+1}\big[\bar{r}_{bt+1} + \bar{P}_{bt+1}(1-\kappa_b)\big],$$

$$\gamma_n g_s \bar{\pi}_t \bar{P}_{st} = \beta \bar{\pi}_{t+1}\big[\bar{r}_{st+1} + \bar{P}_{st+1}(1-\kappa_s)\big],$$

$$\gamma_n g_f \bar{\pi}_t \bar{P}_{ft} = \beta \bar{\pi}_{t+1} \sum_{z_{t+1}} \big[r_{ft+1} + P_{ft+1}(1-\kappa_f)\big] Q(z_t, z_{t+1}),$$

$$p_y(z^t) = \bar{Y}_t^{1-\eta} \bar{y}(z^t)^{\eta-1},$$

$$\bar{r}_l(z^t) = p_y(z^t)(1-\phi) z_t^{(1-\alpha)\phi} \big[\bar{x}(z^t)\big]^{(\lambda-1)/\lambda} \bar{l}_b(z^t)^{-\phi} \bar{k}_b(z^t)^{\alpha\phi} \bar{n}(z^t)^{(1-\alpha)\phi},$$

$$\bar{r}_b(z^t) = p_y(z^t)\alpha\phi z_t^{(1-\alpha)\phi} \big[\bar{x}(z^t)\big]^{(\lambda-1)/\lambda} \bar{l}_b(z^t)^{-\phi} \bar{k}_b(z^t)^{\alpha\phi-1} \bar{n}(z^t)^{(1-\alpha)\phi},$$

$$\bar{\pi}_t \bar{\theta}_t = \psi \ln \frac{\bar{h}(z^t)}{\bar{n}(z^t)} - \psi + \psi \ln g_h^t + \bar{\pi}_t \bar{w}(z^t),$$

$$\bar{n}(z^t) \frac{\psi}{\bar{h}(z^t)} = \pi_t \bar{r}_h(z^t),$$

$$\bar{r}_l(z^t) = (1-\omega)\bar{r}_h(z^t)\bar{l}_h(z^t)^{-\omega} \bar{k}_h(z^t)^{\omega},$$

$$\bar{r}_s(z^t) = \omega \bar{r}_h(z^t)\bar{l}_h(z^t)^{1-\omega} \bar{k}_s(z^t)^{\omega-1},$$

where

$$\bar{w}(z^t) = p_y(z^t)(1-\alpha)\phi z_t^{(1-\alpha)\phi} \big[\bar{x}(z^t)\big]^{(\lambda-1)/\lambda} \times \bar{l}_b(z^t)^{1-\phi} \bar{k}_b(z^t)^{\alpha\phi} \bar{n}(z^t)^{(1-\alpha)\phi-1}$$

and

$$\bar{Y}_t = \Big[\sum q(z^t)\bar{y}(z^t)^{\eta}\Big]^{1/\eta}.$$

It is straightforward to established that the first order conditions for the untransformed economy correspond to these first order conditions once the stationary variables are replaced with their growing counterparts using the transformations given above.

A.3. *Compensation for Lost Growth*

Here we derive the formulas used to evaluate the level increase in consumption and housing required to compensate households for giving up the growth due to local agglomeration:

$$\begin{aligned} g_h &= g_l^{1-\omega} g_s^{\omega} \\ &= \left[\gamma_n^{\zeta-1} g_f^{\zeta}\right]^{1-\omega} g_s^{\omega} \\ &= \left[\gamma_n^{\zeta-1}\left(\frac{g_c}{g_{p_f}}\right)^{\zeta}\right]^{1-\omega}\left(\frac{g_c}{g_{p_s}}\right)^{\omega}. \end{aligned}$$

Therefore, the growth rate of per capita housing absent agglomeration is

$$g_h^* = \left[\hat{\gamma}_n^{\zeta-1}\left(\frac{g_c^*}{\hat{g}_{p_f}}\right)^{\zeta}\right]^{1-\omega}\left(\frac{g_c^*}{\hat{g}_{p_s}}\right)^{\omega}.$$

Notice that the utility of the representative household is

$$\begin{aligned} &\left[\sum_{t=0}^{\infty}\beta^t \ln \bar{C}_t + \sum_{t=0}^{\infty}\beta^t \ln g_c^t + \sum_{t=0}^{\infty}\beta^t \sum q_t(z^t)\bar{n}(z^t)\psi \ln \frac{\bar{h}(z^t)}{\bar{n}(z^t)} \right. \\ &\qquad \left. + \sum_{t=0}^{\infty}\beta^t \sum q_t(z^t)\bar{n}(z^t)\psi \ln g_h^t\right] \\ &= \frac{\ln \bar{C}}{1-\beta} + \frac{\beta \ln g_c}{(1-\beta)^2} + \frac{\ln \dfrac{\bar{h}}{\bar{n}}}{1-\beta} + \psi \frac{\beta \ln g_h}{(1-\beta)^2}. \end{aligned}$$

We seek μ so that utility with and without agglomeration is equated. That is,

$$\begin{aligned} &(1+\psi)\frac{\ln \mu}{1-\beta} + \frac{\ln \bar{C}}{1-\beta} + \frac{\beta \ln g_c^*}{(1-\beta)^2} + \psi \frac{\ln \dfrac{\bar{h}}{\bar{n}}}{1-\beta} + \psi \frac{\beta \ln g_h^*}{(1-\beta)^2} \\ &= \frac{\ln \bar{C}}{1-\beta} + \frac{\beta \ln \hat{g}_c}{(1-\beta)^2} + \psi \frac{\ln \dfrac{\bar{h}}{\bar{n}}}{1-\beta} + \psi \frac{\beta \ln \hat{g}_h}{(1-\beta)^2}. \end{aligned}$$

Solving for μ,

$$\mu = \left(\frac{\hat{g}_c}{g_c^*}\left(\frac{\hat{g}_h}{g_h^*}\right)^{\psi}\right)^{\beta/((1-\beta)(1+\psi))}.$$

APPENDIX B: DATA

This section provides a detailed description of how we construct empirical counterparts to model variables from various data sources and how we merge our different data sources.

B.1. *MSA-Level Panel Data, 1978–2009*

B.1.1. *Current Population Survey Data (Wages and Hours Worked by Skill)*

The March Current Population Survey (CPS) data are available for download at http://cps.ipums.org/cps/ as part of the Integrated Public Use Microdata Series (IPUMS-CPS) project at the University of Minnesota Population Center.

We download the March CPS data from 1979 through 2010. We chose 1979 as our starting year because the number of metropolitan areas we can identify in the CPS and then match to data on housing rents drops off rapidly prior to 1979. The CPS wage and employment questions refer to the "previous calendar year." Therefore, data for any given year's CPS is treated as data appropriate for the previous calendar year. For example, variables generated from the March 2005 CPS would be treated as data for the year 2004.

In each year of our data, we use the following criteria to restrict the sample (with IPUMS-CPS variables shown in italics)

- Respondent lives in a household, not in group quarters or vacant units ($gq = 1$).
- Is aged 20–65 ($age \geq 20$ and $age \leq 65$).
- Wage and salary income in the previous calendar year is identified and is nonzero ($incwage > 0$ and $incwage < 999{,}998$).
- Weeks worked in the previous calendar year is identified and is between 1 and 52 ($wkswork1 \geq 1$ and $weekswork1 \leq 52$).
- Hours worked in a typical week in the previous year (if the respondent worked) is identified and is between 1 and 99 ($uhrswork \geq 1$ and $uhrswork \leq 99$).
- Educational attainment is recorded ($educ \geq 2$ and $educ \leq 115$).
- Has an identified metro area of residence (*metarea* nonmissing).[2]

For each MSA, we use the CPS data to create the following three variables:

(i) Ratio of labor input of high skill to labor input of low skill, $m = n_s/n_u$

(ii) Ratio of total wages paid to total wages paid to low skill workers, s

(iii) Average weekly wage of high skill workers, w_s.

We use the *educ* categorical variable to label respondents as either "low" or "high" skill workers. High skill workers are assumed to have completed 1+ years of college ($educ \geq 80$ and $educ \leq 115$). Everyone else in the sample is assumed to be a low skill worker.

The variable n_s is created as the total of weeks worked the previous calendar year (*wkswork*1) multiplied by the number of hours per week the respondent usually worked (*uhrswork*) for high skill workers. The variable n_u is the

[2]According to notes from the IPUMS-CPS, the metro area of residence was not collected from respondents, but was added by the Census Bureau. The metro areas of residence are based on Federal Information Processing Standards (FIPS) codes used in the 1990 census.

same, but for low skill workers. For each respondent, we weigh the product of *wkswork*1 and *uhrswork* using the IPUMS-CPS sampling person weights, *perwt*.

The variable s is computed as

$$\frac{w_u n_u + w_s n_s}{w_u n_u} = \frac{\sum_{j \in MSA_i} perwt_j \cdot wages_j}{\sum_{j \in MSA_i} perwt_j \cdot wages_j \cdot 1\{unskilled_j\}}$$

for respondent j in MSA i, that is, as the sum of all low and high skill workers' pre-tax wage and salary income for the previous calendar year (*incwage*) divided by the sum of all low skill workers' pre-tax wage and salary income for the previous calendar year. We weigh pre-tax wage and salary income for all persons using the IPUMS-CPS sampling person weights.

The variable w_s is created as the sum of all high skill workers' pre-tax wage and salary income for the previous calendar year (created as an input into s) divided by n_s.

B.1.2. *BEA Data (Output Prices)*

We assume that the price of output varies across MSAs because industry composition varies across MSAs and the price index for industry output varies across industries.

Chain-type price indexes for industry output are available over the 1947–2009 period in the Annual Industry Accounts, http://www.bea.gov/industry/index.htm#annual. To construct a price index for output produced by MSA, we merge this information with MSA-level data on earnings by industry that is available in Tables CA05 and CA05N of the Regional Economic Accounts, http://www.bea.gov/regional/reis/. Earnings is inclusive of wage and salary disbursements, supplements to wages and salaries, and proprietors' income.

In the remainder of this section, Section B.1.2, the notation will differ from that used in the paper.

Denote $g_{t,j}$ as the growth rate of the price of industry output j from periods t to $t+1$, and denote g_t^i as the growth rate of the price of all output produced in MSA i between years t and $t+1$. Assuming output from $j = 1, \dots, N$ industries is produced in MSA i in year t, we set the growth rate of the price of output produced in MSA i between years t and $t+1$ as

$$g_t^i = \sum_{j=1}^{N} \omega_{t,j}^i g_{t,j}. \tag{S23}$$

The weight on each industry, $\omega^i_{t,j}$, is the share of our estimate of total value of the MSA i attributable to value added of industry j in year t:

$$\omega^i_{t,j} = \frac{\mu_j \epsilon^i_{t,j}}{\sum_{k=1}^{N} \mu_j \epsilon^i_{t,k}}, \tag{S24}$$

where $\epsilon^i_{t,j}$ stands for total earnings of employees in industry j in MSA i during year t, and μ_j is a time- and MSA-invariant "markup" that scales earnings of industry j to value added from industry j (described next).[3] For each MSA, we construct a price index for output, normalized to 1.0 in the year 1969, that is consistent with the sequence of time-series estimates of g^i_t.

Before describing how we compute μ_j, we note two details about the earnings and industry data. First, on a somewhat infrequent basis, Tables CA05 and CA05N do not report estimates of earnings for a given industry in an MSA in a given year. In these cases, we set earnings for this industry–MSA–year cell to 0.[4] Also, some of the industry–MSA–year employment estimates are marked with code E. According to the BEA website, these estimates "constitute the major portion of the true estimate." In these cases, we assume that the reported estimate is equal to the actual estimate.

Second, the definition of industries in the Regional Accounts is not consistent across years. Table CA05 reports employment based on SIC industry classifications over the 1969–2000 period and CA05N reports employment based on NAICS industry classifications after 2001.

We map SIC and NAICS industry employment from Tables CA05 and CA05N to prices from the Annual Industry Accounts according to the tables shown later in this section. These tables list all the categories of nonfarm private employment. The sum of the earnings estimates in each of these categories is considered as total nonfarm earnings and is used to compute the denominator of equation (S24).

In all cases except one, there is an exact correspondence of earnings estimates from Tables CA05 and CA05N to prices from the Annual Industry Accounts. For the SIC category of "Transportation and public utilities," line 500 of Table CA05, there is no clean analogous price index in the Annual Industry Accounts. Instead, the Annual Industry Accounts includes separate price indexes for "Transportation and warehousing" and "Utilities." In Table CA05, we therefore separate earnings of the single transportation and public utilities

[3]The markup is allowed to change in 1997, which industry classifications change from Standard Industrial Classification (SIC) based to North American Industry Classification System (NAICS) based.

[4]The three reasons that are listed for omission are (a) to avoid disclosure of confidential information (code D), (b) earnings are less than \$50,000 (code L), or (c) data are not available for this year (code N). These omissions occur in approximately 6% of industry–MSA–year cells from 1969 to the mid-1990s and about 13% of cells after the mid-1990s.

category into earnings in two categories: Earnings from utilities (electric, gas, and sanitary services, line 570) and earnings from transportation and public utilities less earnings from utilities (i.e. line 500 less line 570).

Finally, we need to compute a markup that maps earnings to value added. For each industry, we compute the markup μ_j as the product of two estimated values. The first is the fraction of earnings, by industry, not attributable to proprietor's income. We compute this so as to remove an estimate of proprietors' income from reported earnings by industry by MSA. For each of the SIC industry classifications covering the 1947–1997 period, we compute this fraction using data on the components of value added by industry, available in the file GDPbyInd_VA_SIC, which is available at http://www.bea.gov/industry/io_histannual.htm. Similar data are not available for NAICS, so we map our SIC based estimates to NAICS industries for the 1997–2009 period. Taking the construction industry as an example, in 1947 reported compensation of employees in this industry (in millions) is \$6266 and reported proprietors' income is \$2123. We compute the fraction of earnings not attributable to proprietor's income in this year as $0.747 = 6266/(6266 + 2123)$. We repeat this process each year over the 1947–1997 period and compute the average over all years for the construction sector as 0.788. Thus, for the construction sector, in each MSA in each year we scale reported earnings of construction sector employees by 0.788 to remove an estimate of proprietor's income.

In the second step, we scale the estimate of compensation of employees less proprietors' income to value added. For SIC industries over the 1947–1997 period, we use data from the GDPbyInd_VA_SIC file, and for NAICS industries over the 1998–2009 period, we use similar data from the GDPbyInd_VA_NAICS file, available at http://www.bea.gov/industry/gdpbyind_data.htm, to make this computation. Again, using the example of the construction industry to illustrate how this process works, according to the GDPbyInd_VA_SIC file, in 1947, the reported value added of the industry (in millions) is \$9057 and compensation of employees is \$6266, and thus the ratio of value added to compensation of employees in that year is $1.445 = 9057/6266$. Averaged over all years in the 1947–1997 period, the ratio of value added to compensation of employees in the construction industry is 1.432. We use a similar procedure to compute the mapping of compensation of employees to value added using a similar procedure for the NAICS industries over the 1998–2009 period.

A summary of our procedure for the construction sector is the following: We set value added from the construction sector in MSA i in any year t over 1947–1997 equal to total earnings of employees (from Table CA05) in that year in that MSA multiplied by μ_j for construction, which we compute as $1.128 = 0.788 * 1.432$. We repeat this for every SIC industry (1947–1997) and every NAICS industry (1998–2009) for every MSA in every year. In Tables S.I and S.II, we list our estimates of the two components of μ_j in the right-most columns.

TABLE S.I

EARNINGS AND PRICE DATA BY INDUSTRY, 1969–2001[a]

Data for Earnings, $w^i_{t,j}$ Regional Accounts Table CA05, 1969–2000		Data for Growth in Prices, $g^p_{t,j}$ Industry Accounts, 1969–2001		$\mu_j = a * b$	
				a	b
Line	Label	Line	Label		
100	Agricultural services, forestry fishing, and other	3	Agriculture, forestry, fishing, and hunting	0.300	4.858
200	Mining	6	Mining	0.895	3.092
300	Construction	11	Construction	0.788	1.432
400	Manufacturing	12	Manufacturing	0.979	1.454
500[b]	Transportation and public utilities less electric, gas, and sanitary services	36	Transportation and warehousing	0.932	1.981
570	Electric, gas, and sanitary services	10	Utilities	0.925	3.197
610	Wholesale trade	34	Wholesale trade	0.899	1.873
620	Retail trade	35	Retail trade	0.806	1.721
700	Finance, insurance, and real estate	50	Finance, insurance, real estate, rental, and leasing	0.856	4.784
800	Services	59	Professional and business services	0.742	1.557
900	Government and government enterprises	82	Government	1.000	1.236

[a] a: Adjustment to remove proprietor's income from earnings. b: Mapping of wage compensation to value added.

[b] See text for details.

B.1.3. *BLS Data and 1990 Decennial Census of Housing (Housing Rents)*

We create annual estimates over the 1978–2009 period of the average rents paid for certain types of rental units, by MSA, using a two-step procedure.

In the first step, we estimate the average rents paid for certain types of rental housing units in 1990 using household-level data from the 1990 Decennial Census of Housing (DCH). These data are available for download at http://usa.ipums.org/usa/ as part of the Integrated Public Use Microdata Series (IPUMS-USA) project at the University of Minnesota Population Center. We use data from the 1990 DCH.

With IPUMS-USA variables in italics, we restrict the 1990 DCH sample to renter nonfarm households in 2–19 unit residences in a building built between 1940 and 1986, and living in an identifiable MSA ($ownershg = 2$, $farm \neq 1$, $unitsstr \in \{5, 8\}$, $builtyr \in \{3, 7\}$, and $metarea > 0$) who live in households and do not live in group quarters ($gq \in \{3, 4, 6\}$), and where the reported monthly gross rent of the house (rent inclusive of utilities) is nonzero ($rentgrs > 0$). Conditional on these restrictions, we compute the weighted average value of units by MSA using the sampling weight variable *hhwt*. These calculations yield estimates of the average rental price of housing for 272 metro areas as identified in the 1990 DCH. We exclude single-family rented units, rented high-rise units

TABLE S.II

EARNINGS AND PRICE DATA BY INDUSTRY, 2001–2006[a]

Data for Earnings, $w^i_{t,j}$ Regional Accounts Table CA05N, 2001–2005		Data for Growth in Prices, $g^p_{t,j}$ Industry Accounts, 2001–2006		$\mu_j = a * b$	
Line	Label	Line	Label	a	b
100	Forestry, fishing, related activities, and other	5	Forestry, fishing, and related activities	0.300	3.441
200	Mining	6	Mining	0.895	3.397
300	Utilities	10	Utilities	0.925	3.737
400	Construction	11	Construction	0.788	1.520
500	Manufacturing	12	Manufacturing	0.979	1.659
600	Wholesale trade	34	Wholesale trade	0.899	1.889
700	Retail trade	35	Retail trade	0.806	1.729
800	Transportation and warehousing	36	Transportation and warehousing	0.932	1.541
900	Information	45	Information	0.742	2.203
1000	Finance and insurance	51	Finance and insurance	0.856	1.888
1100	Real estate and rental and leasing	56	Real estate and rental and leasing	0.856	15.856
1200	Professional, scientific, and technical services	60	Professional, scientific, and technical services	0.742	1.534
1300	Management of companies and enterprises	64	Management of companies and enterprises	0.742	1.174
1400	Administrative and waste services	65	Administrative and waste management services	0.742	1.329
1500	Educational services	69	Educational services	0.742	1.132
1600	Health care and social assistance	70	Health care and social assistance	0.742	1.218
1700	Arts, entertainment, and recreation	75	Arts, entertainment, and recreation	0.742	1.705
1800	Accommodation and food services	78	Accommodation and food services	0.742	1.625
1900	Other services except public administration	81	Other services except government	0.742	1.546
2000	Government and government enterprises	82	Government	1.000	1.236

[a] a: Adjustment to remove proprietor's income from earnings. b: Mapping of wage compensation to value added.

(> 20 units), and units in very old (built before 1940) or very new (built after 1986) apartment buildings to attempt to keep the average characteristics of rental units roughly constant across metropolitan areas without resorting to hedonic regressions.

In the second step, we extrapolate the annual rental price of housing in each metro area forward from 1990 to 2009 and backward from 1990 to 1978 using annual MSA-specific constant-quality price indexes for the price per unit of shelter. These price indexes for shelter are published by the Bureau of Labor Statistics (BLS) as part of computations for the Consumer Price Index and are available at http://www.bls.gov. The BLS reports rental price indexes for

27 MSAs, but the indexes of three of these MSAs (Phoenix, AZ, Washington, DC, and Tampa Bay, FL) do not have data available prior to 1985, so we exclude these from our sample. The CPS does not have data on Anchorage and Honolulu back to 1978, explaining our sample of 22 MSAs.

In 1983, the BEA changed its procedure for measuring the price of owner-occupied rent, which accounts for about 73% of all spending on shelter. After 1983, the BEA began measuring the price of owner-occupied rent using the "rental equivalence" approach, whereas in earlier years, the BLS used the "asset price method."[5] To eliminate this nontrivial inconsistency in the data, we replace the reported values of the shelter indexes from 1978–1982 with predicted values, essentially predicting what the BLS would have reported if the owner-occupied data had been collected using the "rental equivalence approach." Specifically, we regress the log BLS shelter indexes on MSA dummies and the log BLS tenant-rent indexes over the 1983–2009 period. The R^2 of the regression is 0.99 and the coefficient on log tenant rents is 1.055. Based on the regression results and the values of the log tenant-rent indexes in the 1978–1982 period, we predict the log MSA shelter indexes from 1978–1982.

B.1.4. *Merging the MSA-Level Data, 1978–2009*

We merge the CPS data on wages and employment (Section B.1.1) with the BEA data on output prices (Section B.1.2) and the annual data we construct on housing rents (Section B.1.3). The data are merged by MSA and by year. After all data are merged, we are left with a balanced panel of 22 MSAs. In every MSA and date in our sample, the minimum number of respondents from the CPS is never less than 200; it is typically about 250 until 1999 and then jumps to about 450 after 2000. The median number of respondents is about 540 until about 2000, at which point the median jumps to about 1000. The maximum number of respondents is always above 3000 and is typically about 4000.

The MSAs are defined as sets of counties, but MSA definitions are not completely consistent across data sources or across time.

- The BEA fixes MSA definitions: The counties that comprise MSAs are identical in every year of reported data. As of the writing of this paper, the MSA definitions in the BEA data are given by the list in the December 2009 report of the Office of Management and Budget (OMB).[6] If the OMB update MSA definitions, the BEA data revise history accordingly.
- The definitions of MSAs change over time in the CPS data and the BLS rental price index data. To our knowledge, the MSA definitions over history do not revise in either of these data sets. The BLS has three different sets of MSA definitions: 1978–1986, 1987–1997, and 1998 onward. The CPS has four

[5]See http://www.bls.gov/cpi/cpifact6.htm for details.
[6]See http://www.census.gov/population/www/metroareas/metrodef.html.

TABLE S.III

MSA DEFINITIONS BY YEAR, ATLANTA AND BOSTON

	Atlanta		Year Added to Sample			Boston		Year Added to Sample	
FIPS	County	State	CPS	BLS	FIPS	County	State	CPS	BLS
13035	Butts	GA	1978*	1978	25009	Essex (pt.)	MA	1978	1978
13057	Cherokee	GA	1978	1978	25017	Middlesex (pt.)	MA	1978	1978
13063	Clayton	GA	1978	1978	25021	Norfolk (pt.)	MA	1978	1978
13067	Cobb	GA	1978	1978	25023	Plymouth (pt.)	MA	1978	1978
13089	DeKalb	GA	1978	1978	25025	Suffolk	MA	1978	1978
13097	Douglas	GA	1978	1978	25005	Bristol (pt.)	MA	1985	1987
13113	Fayette	GA	1978	1978	33011	Hillsborough (pt.)	NH	1985	1987
13117	Forsyth	GA	1978	1978	33015	Rockingham (pt.)	NH	1985	1987
13121	Fulton	GA	1978	1978	25027	Worcester (pt.)	MA	1985	1987
13135	Gwinnett	GA	1978	1978	25013	Hampden (pt.)	CT	1994	1998
13151	Henry	GA	1978	1978	33013	Merrimack (pt.)	NH	1994	1998
13217	Newton	GA	1978	1978	33017	Strafford (pt.)	NH	1994	1998
13223	Paulding	GA	1978	1978	09015	Windham (pt.)	CT	1994	1998
13247	Rockdale	GA	1978	1978	23031	York (pt.)	ME	1994	1998
13297	Walton	GA	1978	1978					
13013	Barrow	GA	1985	1987					
13077	Coweta	GA	1985	1987					
13255	Spalding	GA	1985	1987					
13015	Bartow	GA	1994	1998					
13045	Carroll	GA	1994	1998					
13227	Pickens	GA	1994	1998					
13085	Dawson	GA	2004	x					
13143	Haralson	GA	2004	x					
13149	Heard	GA	2004	x					
13159	Jasper	GA	2004	x					
13171	Lamar	GA	2004	x					
13199	Meriwether	GA	2004	x					
13231	Pike	GA	2004	x					

*1994 deleted, 2004 added.

different sets of MSA definitions: 1978–1984, 1985–1993, 1994–2003, and 2004 onward.[7]

Tables S.III–S.XIII list the counties that comprise each MSA and the year in which the counties were added to the CPS and the BLS MSA-level data. Sometimes a county is included in the CPS definition but not the BLS definition. We indicate these cases with the letter "x." Information used to construct the list of counties in the MSA definitions in the CPS by year are available on

[7]We lag the reported CPS date by one year for reasons discussed earlier.

TABLE S.IV

MSA DEFINITIONS BY YEAR, CHICAGO AND CINCINNATI

	Chicago		Year Added to Sample			Cincinnati		Year Added to Sample	
FIPS	County	State	CPS	BLS	FIPS	County	State	CPS	BLS
17031	Cook	IL	1978	1978	39025	Clermont	OH	1978	1978
17043	DuPage	IL	1978	1978	39061	Hamilton	OH	1978	1978
17089	Kane	IL	1978	1978	39165	Warren	OH	1978	1978
18089	Lake	IL	1978	1978	21015	Boone	KY	1978	1978
17111	McHenry	IL	1978	1978	21037	Campbell	KY	1978	1978
17197	Will	IL	1978	1978	21117	Kenton	KY	1978	1978
18089	Lake	IN	1985	1978	18029	Dearborn	IN	1978	1978
18127	Porter	IN	1985	1978	39017	Butler	OH	1985	1987
17063	Grundy	IL	1985	1987	18115	Ohio	IN	1994	1998
17093	Kendall	IL	1985	1987	21077	Gallatin	KY	1994	1998
55059	Kenosha	WI	1985	1987	21081	Grant	KY	1994	1998
17037	DeKalb	IL	1994	1998	21191	Pendleton	KY	1994	1998
17091	Kankakee	IL	1994	1998	39015	Brown	OH	1994	1998
18073	Jasper	IN	2004	x	21023	Bracken	KY	2004	x
18111	Newton	IN	2004	x	18047	Franklin	IN	2004	x

TABLE S.V

MSA DEFINITIONS BY YEAR, DALLAS AND DENVER

	Dallas		Year Added to Sample			Denver		Year Added to Sample	
FIPS	County	State	CPS	BLS	FIPS	County	State	CPS	BLS
48085	Collin	TX	1978	1978	08001	Adams	CO	1978	1978
48113	Dallas	TX	1978	1978	08005	Arapahoe	CO	1978	1978
48121	Denton	TX	1978	1978	08013	Boulder	CO	1978*	1978
48139	Ellis	TX	1978	1978**	08031	Denver	CO	1978	1978
48221	Hood	TX	1978*	1978	08047	Gilpin	CO	1978++	1978***
48251	Johnson	TX	1978	1978	08059	Jefferson	CO	1978	1978
48257	Kaufman	TX	1978	1978	08035	Douglas	CO	1978	1987
48367	Parker	TX	1978	1978	08123	Weld	CO	1994+	1998
48397	Rockwall	TX	1978	1978	08014	Broomfield	CO	2004	x
48439	Tarrant	TX	1978	1978	08019	Clear Creek	CO	2004	x
48497	Wise	TX	1978	1978***	08039	Elbert	CO	2004	x
48213	Henderson	TX	1994	1998	08093	Park	CO	2004	x
48231	Hunt	TX	1994	1998					
48119	Delta	TX	2004	x					

*1985 deleted, 1994 added, 2004 deleted; **1987 deleted, 1998 added; ***1987 deleted, +2004 deleted, ++1985 deleted, 2004 added.

TABLE S.VI

MSA DEFINITIONS BY YEAR, DETROIT AND HOUSTON

	Detroit		Year Added to Sample			Houston		Year Added to Sample	
FIPS	County	State	CPS	BLS	FIPS	County	State	CPS	BLS
26087	Lapeer	MI	1978	1978	48039	Brazoria	TX	1978	1978
26093	Livingston	MI	1978	1978	48157	Fort Bend	TX	1978	1978
26099	Macomb	MI	1978	1978	48201	Harris	TX	1978	1978
26125	Oakland	MI	1978	1978	48291	Liberty	TX	1978	1978
26147	St. Clair	MI	1978	1978	48339	Montgomery	TX	1978	1978
26163	Wayne	MI	1978	1978	48473	Waller	TX	1978	1978
26161	Washtenaw	MI	1985	1987	48167	Galveston	TX	1985	1987
26115	Monroe	MI	1985	1998	48071	Chambers	TX	1994	1998
26049	Genesee	MI	1994	1998	48015	Austin	TX	2004	x
26091	Lenawee	MI	1994	1998	48407	San Jacinto	TX	2004	x

TABLE S.VII

MSA DEFINITIONS BY YEAR, KANSAS CITY AND LOS ANGELES

	Kansas City		Year Added to Sample			Los Angeles		Year Added to Sample	
FIPS	County	State	CPS	BLS	FIPS	County	State	CPS	BLS
29037	Cass	MO	1978	1978	06037	Los Angeles	CA	1978	1978
29047	Clay	MO	1978	1978	06059	Orange	CA	1978	1978
29095	Jackson	MO	1978	1978	06065	Riverside	CA	1985*	1987
29165	Platte	MO	1978	1978	06071	San Bernardino	CA	1985*	1987
29177	Ray	MO	1978	1978	06111	Ventura	CA	1985*	1987
20091	Johnson	KS	1978	1978					
20209	Wyandotte	KS	1978	1978					
29107	Lafayette	MO	1985	1987					
20103	Leavenworth	KS	1985	1987					
20121	Miami	KS	1985	1987					
29049	Clinton	MO	1994	1998					
29013	Bates	MO	2004	x					
29025	Caldwell	MO	2004	x					
20049	Franklin	KS	2004	x					
20107	Linn	KS	2004	x					

*2004 deleted.

TABLE S.VIII

MSA DEFINITIONS BY YEAR, MIAMI AND MILWAUKEE

	Miami		Year Added to Sample			Milwaukee		Year Added to Sample	
FIPS	County	State	CPS	BLS	FIPS	County	State	CPS	BLS
12086	Dade	FL	1978	1978	55079	Milwaukee	WI	1978	1978
12011	Broward	FL	1985	1987	55089	Ozaukee	WI	1978	1978
12099	Palm Beach	FL	2004	x	55131	Washington	WI	1978	1978
					55133	Waukesha	WI	1978	1978
					55101	Racine	WI	1985	1998

the NBER and the BLS website.[8] Information used to construct the counties in the MSA definitions in the BLS for the 1978–1996 and 1987–1997 time periods were supplied to us by the BLS and are available online for 1998 onward.[9]

To assess how the changing nature of MSA definitions may affect our empirical work, for the years 1980, 1990, 2000, and 2008, we compute the population in each MSA according to (a) the 1978 BLS definitions of MSAs and (b) the 1998 BLS definitions. These computations are shown in Table S.XIV. In most MSAs, 80% or more of the population lives in the counties included as part of the 1978 BLS definitions. In every MSA, the ratio is at least 60%. In addition, the ratio of the population defined according to the 1978 definition to the population defined according to the 1998 definition is stable in almost every MSA. This suggests that the population of most counties included in the 1998 MSA definitions increased at roughly the same rate over the 1980–2008 period.

B.2. *Aggregate Data*

B.2.1. *Data Used for the Depreciation Rate of Residential Structures*

One of our moment conditions that involves the depreciation rate on residential structures, κ_h, is

$$E\left[\kappa_h - \frac{P_{ht}D_{ht}}{P_{ht}K_{ht}}\right] = 0,$$

where $P_{ht}D_{ht}$ is nominal value of aggregate depreciation on structures in year t and $P_{ht}K_{ht}$ is the nominal value of the aggregate stock of structures in year t. Our data on $P_{ht}D_{ht}$ are from line 7 (Residential Fixed Assets) of the BEA Fixed Assets Table 1.3, Current Cost Depreciation of Fixed Assets and Consumer

[8]See http://www.nber.org/cps-basic/metrochg.pdf and point 6 of http://www.bls.gov/gps/notescps.htm.

[9]See Appendix 5 of http://www.bls.gov/opub/hom/pdf/homch17.pdf.

TABLE S.IX
MSA DEFINITIONS BY YEAR, MINNEAPOLIS AND NEW YORK

	Minneapolis		Year Added to Sample			New York		Year Added to Sample	
FIPS	County	State	CPS	BLS	FIPS	County	State	CPS	BLS
27003	Anoka	MN	1978	1978	34003	Bergen	NJ	1978	1978
27019	Carver	MN	1978	1978	36005	Bronx	NY	1978	1978
27025	Chisago	MN	1978	1978	36047	Kings	NY	1978	1978
27037	Dakota	MN	1978	1978	36061	New York	NY	1978	1978
27053	Hennepin	MN	1978	1978	36079	Putnam	NY	1978	1978
27123	Ramsey	MN	1978	1978	36081	Queens	NY	1978	1978
27139	Scott	MN	1978	1978	36085	Richmond	NY	1978	1978
27163	Washington	MN	1978	1978	36087	Rockland	NY	1978	1978
27171	Wright	MN	1978	1978	36119	Westchester	NY	1978	1978
55109	St. Croix	WI	1978	1978	36059	Nassau	NY	1978	1978
27059	Isanti	MN	1985	1987	36103	Suffolk	NY	1978	1978
27141	Sherburne	MN	1994	1998	34013	Essex	NJ	1978	1978
55093	Pierce	WI	1994	1998	34027	Morris	NJ	1978	1978
					34039	Union	NJ	1978	1978
					34035	Somerset	NJ	1978	1978
					34017	Hudson	NJ	1978	1978
					34023	Middlesex	NJ	1978	1978
					34031	Passaic	NJ	1978	1978
					9001	Fairfield (pt.)	CT	1985*	1987
					9005	Litchfield (pt.)	CT	1985*	1987
					9009	New Haven (pt.)	CT	1985*	1987
					36071	Orange	NY	1985*	1987
					34019	Hunterdon	NJ	1985	1987
					34025	Monmouth	NJ	1985	1987
					34029	Ocean	NJ	1985	1987
					34037	Sussex	NJ	1985	1987
					36027	Dutchess	NY	1985*	1998
					9007	Middlesex (pt.)	CT	1985*	1998
					34021	Mercer	NJ	1985*	1998
					34041	Warren	NJ	1985*	1998
					42103	Pike	PA	1994	1998

*2004 deleted.

Durable Goods.[10] Our data on $P_{ht}K_{ht}$ are from line 7 of the BEA Fixed Assets Table 1.1, Current Cost Net Stock of Fixed Assets and Consumer Durable Goods. The capital stocks reported in Fixed Assets Table 1.1 are year-end values. To adjust for this, we set K_{ht} as the once-lagged reported year-end value, that is, we set K_{ht} for the year 2000 as the year-end reported value for 1999.

[10]These tables are available at http://www.bea.gov/national/FA2004/SelectTable.asp.

TABLE S.X

MSA DEFINITIONS BY YEAR, PHILADELPHIA AND PITTSBURGH

	Philadelphia		Year Added to Sample			Pittsburgh		Year Added to Sample	
FIPS	County	State	CPS	BLS	FIPS	County	State	CPS	BLS
42017	Bucks	PA	1978	1978	41994	Allegheny	PA	1978	1978
42029	Chester	PA	1978	1978	42007	Beaver	PA	1978	1978
42045	Delaware	PA	1978	1978	42125	Washington	PA	1978	1978
42091	Montgomery	PA	1978	1978	42129	Westmoreland	PA	1978	1978
42101	Philadelphia	PA	1978	1978	42051	Fayette	PA	1985	1987
34005	Burlington	NJ	1978	1978	42019	Butler	PA	1994	1998
34007	Camden	NJ	1978	1978	42004	Armstrong	PA	2004	x
34015	Gloucester	NJ	1978	1978					
10003	New Castle	DE	1992	1987					
24015	Cecil	MD	1992	1987					
34011	Cumberland	NJ	1992	1987					
34021	Mercer	NJ	1992	1987					
34033	Salem	NJ	1992	1987					
34001	Atlantic	NJ	1998	1998					
34009	Cape May	NJ	1998	1998					

B.2.2. *Data Used for the Depreciation Rate of Infrastructure Capital*

One of our moment conditions that involves the depreciation rate on residential structures, κ_f, is

$$E\left[\kappa_f - \frac{P_{ft}D_{ft}}{P_{ft}K_{ft}}\right] = 0,$$

TABLE S.XI

MSA DEFINITIONS BY YEAR, PORTLAND AND SAN DIEGO

	Portland		Year Added to Sample			San Diego		Year Added to Sample	
FIPS	County	State	CPS	BLS	FIPS	County	State	CPS	BLS
41005	Clackamas	OR	1978	1978	06073	San Diego	CA	1978	1978
41051	Multnomah	OR	1978	1978					
41067	Washington	OR	1978	1978					
53011	Clark	WA	1978	1978					
41071	Yamhill	OR	1985	1987					
41009	Columbia	OR	1994	1998					
41047	Marion	OR	1994*	1998					
41053	Polk	OR	1994*	1998					
53059	Skamania	WA	2004	x					

*2004 deleted.

TABLE S.XII

MSA DEFINITIONS BY YEAR, SAN FRANCISCO AND SEATTLE

	San Francisco		Year Added to Sample			Seattle		Year Added to Sample	
FIPS	County	State	CPS	BLS	FIPS	County	State	CPS	BLS
06001	Alameda	CA	1978	1978	53033	King	WA	1978	1978
06013	Contra Costa	CA	1978	1978	53061	Snohomish	WA	1978	1978
06041	Marin	CA	1978	1978	53053	Pierce	WA	1985	1987
06075	San Francisco	CA	1978	1978	53029	Island	WA	1994	1998
06081	San Mateo	CA	1978	1978	53035	Kitsap	WA	1994	1998
06055	Napa	CA	1985*	1987	53067	Thurston	WA	1994	1998
06085	Santa Clara	CA	1985*	1987					
04023	Santa Cruz	CA	1985*	1987					
06095	Solano	CA	1985*	1987					
06097	Sonoma	CA	1985*	1987					

*2004 deleted.

where $P_{ft}D_{ft}$ is the nominal value of aggregate depreciation on infrastructure capital in year t and $P_{ft}K_{ft}$ is the nominal value of the aggregate stock of infrastructure in year t.

TABLE S.XIII

MSA DEFINITIONS BY YEAR, ST. LOUIS AND TAMPA BAY

	St. Louis		Year Added to Sample			Tampa Bay		Year Added to Sample	
FIPS	County	State	CPS	BLS	FIPS	County	State	CPS	BLS
29071	Franklin	MO	1978	1978	12057	Hillsborough	FL	1978	1978
29099	Jefferson	MO	1978	1978	12101	Pasco	FL	1978	1978
29183	St. Charles	MO	1978	1978	12103	Pinellas	FL	1978	1978
29189	St. Louis	MO	1978	1978	12053	Hernando	FL	1985	1987
29510	St. Louis City	MO	1978	1978					
17027	Clinton	IL	1978	1978					
17119	Madison	IL	1978	1978					
17133	Monroe	IL	1978	1978					
17163	St. Clair	IL	1978	1978					
17083	Jersey	IL	1985	1987					
29113	Lincoln	MO	1994	1998					
29219	Warren	MO	1994	1998					
29055	Crawford (pt.)	MO	1994	x					
29221	Washington	MO	2004	x					
17005	Bond	IL	2004	x					
17013	Calhoun	IL	2004	x					
17117	Macoupin	Il	2004	x					

TABLE S.XIV

POPULATION (IN MILLIONS) BY MSA DEFINITION (1978 VERSUS 1998)

		1980	1990	2000	2008		1980	1990	2000	2008
1978 defn	Atlanta	2.03	2.70	3.75	4.75	Milwaukee	1.40	1.43	1.50	1.55
1998 defn		2.25	2.98	4.13	5.24		1.57	1.61	1.69	1.75
1978/1998		0.90	0.91	0.91	0.91		0.89	0.89	0.89	0.89
1978 defn	Boston	3.66	3.78	4.00	4.10	Minneapolis	2.11	2.44	2.84	3.06
1998 defn		6.11	6.53	6.95	7.18		2.20	2.54	2.97	3.23
1978/1998		0.60	0.58	0.58	0.57		0.96	0.96	0.96	0.95
1978 defn	Chicago	7.31	7.35	8.08	8.39	New York	15.29	15.59	16.88	17.46
1998 defn		7.67	7.72	8.51	8.92		18.81	19.45	21.09	21.88
1978/1998		0.95	0.95	0.95	0.94		0.81	0.80	0.80	0.80
1978 defn	Cincinnati	1.40	1.45	1.55	1.66	Philadelphia	4.72	4.86	5.04	5.14
1998 defn		1.73	1.82	1.98	2.12		5.96	6.22	6.54	6.73
1978/1998		0.81	0.80	0.79	0.78		0.79	0.78	0.77	0.76
1978 defn	Dallas	2.97	3.95	5.12	6.26	Pittsburgh	0.81	0.76	0.75	0.74
1998 defn		3.07	4.07	5.27	6.42		1.12	1.06	1.08	1.07
1978/1998		0.97	0.97	0.97	0.97		0.73	0.72	0.70	0.69
1978 defn	Denver	1.22	1.35	1.70	1.88	Portland	1.24	1.41	1.79	2.05
1998 defn		1.74	1.98	2.58	2.94		1.58	1.79	2.27	2.59
1978/1998		0.70	0.68	0.66	0.64		0.78	0.79	0.79	0.79
1978 defn	Detroit	4.35	4.25	4.45	4.43	St. Louis	2.36	2.42	2.52	2.61
1998 defn		5.29	5.19	5.46	5.46		2.41	2.49	2.60	2.72
1978/1998		0.82	0.82	0.82	0.81		0.98	0.97	0.97	0.96
1978 defn	Houston	2.91	3.49	4.39	5.36	San Diego	1.86	2.50	2.81	3.00
1998 defn		3.12	3.73	4.67	5.68		1.86	2.50	2.81	3.00
1978/1998		0.93	0.94	0.94	0.94		1.00	1.00	1.00	1.00
1978 defn	Kansas City	1.33	1.45	1.63	1.78	San Francisco	3.25	3.69	4.12	4.27
1998 defn		1.45	1.58	1.78	1.94		5.20	6.05	6.82	7.09
1978/1998		0.92	0.91	0.92	0.92		0.63	0.61	0.60	0.60
1978 defn	Los Angeles	9.41	11.27	12.37	12.87	Seattle	1.61	1.97	2.34	2.56
1998 defn		11.50	14.53	16.37	17.79		2.41	2.97	3.55	3.91
1978/1998		0.82	0.78	0.76	0.72		0.67	0.66	0.66	0.65
1978 defn	Miami	1.63	1.94	2.25	2.40	Tampa Bay	1.57	1.97	2.27	2.56
1998 defn		2.64	3.19	3.88	4.15		1.61	2.07	2.40	2.73
1978/1998		0.61	0.61	0.58	0.58		0.97	0.95	0.95	0.94

- We compute the nominal value of infrastructure capital, $P_{ft}K_{ft}$, as the sum of the nominal stocks of (a) federal nondefense, and state and local government highways and streets, (b) federal nondefense, and state and local "other structures" (pre-1996) and transportation and power structures (post-1997), (c) state and local sewer systems structures, (d) state and local water supply facilities, and (e) privately owned power and communication, transportation, and "other" structures. The data for the nominal stocks of federal and state and local infrastructure capital are in the BEA Fixed Asset Tables 7.1a (lines 38, 40, 49, 51, 52, and 53, covering the period 1925–1996) and 7.1b (lines 41, 42, 43, 56, 57, 58, 59, and 60, covering 1997–2009). The data for the nominal stocks of private infrastructure capital are in the BEA Fixed Asset Table 2.1, lines 50, 63, and 67.[11]

- We compute the nominal value of the depreciation of infrastructure capital, $P_{ft}D_{ft}$, analogously. The data for the nominal depreciation of infrastructure capital for the federal and state and local government are in BEA Fixed Asset Tables 7.3a (lines 38, 40, 49, 51, 52, and 53) and 7.3b (lines 41, 42, 43, 56, 57, 58, 59, and 60). Depreciation for privately owned infrastructure capital is reported in BEA Fixed Asset Table 2.4, lines 50, 63, and 67.

As mentioned earlier, the BEA reports the capital stocks data at year end. To adjust for this, we define K_{ft} as the lag of reported year-end values.

B.2.3. *Data Used for the Depreciation Rate of Business Capital*

One of our moment conditions that involves the depreciation rate on capital used in production, κ_b, is

$$E\left[\kappa_b - \frac{P_{bt}D_{bt}}{P_{bt}K_{bt}}\right] = 0,$$

where $P_{bt}D_{bt}$ is the nominal value of aggregate depreciation on capital used in production in year t and $P_{bt}K_{bt}$ is the nominal aggregate stock of capital used in production in year t.

- We compute $P_{bt}D_{bt}$ as the nominal depreciation of all fixed assets and consumer durable goods (line 1) less nominal depreciation of private residential structures (line 7) of the BEA Fixed Assets Table 1.3, all less nominal depreciation of infrastructure capital, defined in Section B.2.2.
- We compute $P_{bt}K_{bt}$ as the nominal stock of all fixed assets and consumer durable goods (line 1) less the nominal stock of private residential structures (line 7) of the BEA Fixed Assets Table 1.1, all less nominal depreciation of infrastructure capital as defined in Section B.2.2.

As mentioned earlier, the BEA reports the capital stocks data at year end, and we adjust for this by setting $P_{bt}K_{bt}$ equal to the lag of reported year-end values.

B.2.4. *Data Used for the Growth Rate of the Price of Housing Structures, Infrastructure Capital, and Business Capital*

Our moment conditions that involve the trend growth rate of the aggregate real price of housing structures g_{PS} and the trend growth rate of the aggregate

[11]The BEA notes that "other" government structures consist "primarily of electric and gas facilities, transit systems, and airfields," whereas "other" private structures include structures that pertain to "water supply, sewage and waste disposal, public safety, highway and street, and conservation and development."

real price of business capital g_{p_b} are

$$E\{(\ln P_{ht} - \ln(g_{p_h})t)t\} = 0,$$
$$E\{(\ln P_{ft} - \ln(g_{p_f})t)t\} = 0,$$
$$E\{(\ln P_{bt} - \ln(g_{p_b})t)t\} = 0.$$

- The variable P_{ht} is the real price for housing structures, defined as the nominal price index for structures divided by the price index of consumption. The nominal price index for structures is computed as the nominal stock of housing structures, line 7 of BEA Fixed Asset Table 1.1, divided by the chain-type quantity index for residential structures, line 7 of BEA Fixed Asset Table 1.2.
- The variable P_{ft} is the real price for infrastructure capital, defined as the nominal price index for infrastructure capital divided by the price index of consumption. The nominal price index for infrastructure capital is computed by chain-weighting the price indexes of each of the components of infrastructure capital described in Section B.2.2: federal nondefense, and state and local government highways and streets, and other (pre-1996) or transportation and power (post-1997), state and local sewer systems structures, state and local water supply facilities, state and local transportation structures and power structures, and privately owned power and communication structures, transportation structures, and other structures. The price indexes for each of the components is computed as the ratio of the nominal stock (Fixed Asset Tables 7.1a, 7.1b, and 2.1) to the chain-type quantity indexes (Fixed Asset Tables 7.2a, 7.2b, and 2.2).
- The variable P_{bt} is the real price for business capital, defined as the nominal price index for business capital divided by the price index of consumption. We compute the nominal price index for business capital by chain-weighting the price index for (a) all fixed assets and consumer durable goods less (b) the price index for housing structures less (c) the price index for infrastructure capital.
  – The nominal price index for all fixed assets and consumer durable goods is computed by dividing the nominal stock of all fixed assets and consumer durable goods, line 1 of BEA Fixed Asset Table 1.1, by the chain-type quantity index for all fixed assets and consumer durable goods, line 1 of BEA Fixed Asset Table 1.2.
  – The price indexes for housing structures and infrastructure capital are defined above.
- We discuss how we create the price index for consumption in Section B.2.9.

B.2.5. *Data Used for the Growth Rate of Housing, Structures, and Land Prices*

Our moment condition that involves the average of rental prices across MSAs, the aggregate growth rate of the real price of land rents g_{p_l} and the real price of housing structures g_{p_h}, and the share of housing rents attributable to housing structures ω is

$$E\{(\ln E_t r_{hit} - [(1-\omega)\ln(g_{p_l}) + \omega \ln(g_{p_h})]t)t\} = 0.$$

We compute $E_t r_{hit}$ in each period as the average level of real rental prices in each MSA. In each MSA, the real rental price is computed as the nominal rental price divided by the price index for consumption. We discuss how we create the price index for consumption in Section B.2.9.

B.2.6. *Data Used for the Structures' Share of Housing Rents*

One of our moment conditions for the structures' share of housing rents, ω, the growth rate of the price of land rents, g_{p_l}, the growth rate of the price of housing structures, g_{p_h}, and the depreciation rate on housing structures, κ_s, is

$$E\left(\frac{\sum p_{lit} l_{hit}}{\sum(P_{ht} k_{hit} + p_{lit} l_{hit})}\left[\frac{\omega}{1-\omega}\frac{R/g_{p_l} - (1-\kappa_f)^{\zeta}}{R/g_{p_h} + \kappa_h - 1} + 1\right] - 1\right) = 0.$$

We set $\sum p_{lit} l_{hit}$ as the market value aggregate value of finished land in residential use, taken from a study by Davis and Heathcote (2007) and available at http://www.lincolninst.edu/subcenters/land-values/price-and-quantity.asp. We compute the annual data as the average of the reported quarterly data. We set $\sum(P_{ht} k_{hit} + p_{lit} l_{hit})$ as the market value of housing (land and structures), taken from the same Davis and Heathcote (2007) study. Again, we set annual values as the average of the reported quarterly values.

B.2.7. *Parameters of the Production Function Related to Capital and Land Shares of Production*

Two of our moment conditions related to capital's and finished land's share of production, α and ϕ, are

$$E\left(\frac{\sum p_{lit} l_{bit}}{\sum(P_{bt} k_{bit} + p_{lit} l_{bit})}\left[\frac{\alpha\phi}{1-\phi}\frac{R/g_{p_l} - (1-\kappa_f)^{\zeta}}{R/g_{p_b} + \kappa_b - 1} + 1\right] - 1\right) = 0, \tag{S25}$$

$$E\left(\frac{\sum w_{it} n_{it}}{\sum[w_{it} n_{it} + r_{lit} l_{bit} + r_{bt} k_{bit}]} - \phi(1-\alpha)\right) = 0. \tag{S26}$$

Data for Equation (S25). We set $\sum p_{lit} l_{bit}$ equal to the aggregate value of finished land used in production, computed as the sum of the following variables:

(i) The value of land used for nonresidential purposes by nonfarm nonfinancial corporate businesses. These data come from Table B.102 of the Flow of Funds Accounts of the United States (FFA). We set the value of land equal to the value of real estate owned by this sector (line 3) less the replacement cost of structures owned by this sector (lines 33 and 34). We set the annual as the average of the reported quarterly observations.

(ii) The value of land used for nonresidential purposes by nonfarm nonfinancial noncorporate businesses. These data come from Table B.103 of the FFA. We set the value of land equal to the value of nonresidential real estate owned by this sector (line 5) less the replacement cost of nonresidential structures owned by this sector (line 33). We set the annual as the average of the reported quarterly observations.

(iii) The current cost of privately owned infrastructure capital (power and communication structures, transportation structures, and other structures) as described in Section B.2.2.

(iv) The value of land used for nonresidential purposes by financial corporations. We compute this as $\mathcal{R}$ (to be defined later) times the Current Cost Net Stock of Private Nonresidential Fixed Assets Owned by Financial Corporations, line 25 of BEA Fixed Asset Table 4.1. We set the annual as the average of the current and lagged year-end observations.

(v) The value of land used for nonresidential purposes by nonprofit organizations. We compute this as $\mathcal{R}$ times the sum of equipment and software owned by nonprofit organizations (line 6 of FFA Table B.100) and the replacement cost of nonresidential structures owned by nonprofit organizations (line 46 of FFA Table B.100). We set the annual as the average of the reported quarterly observations.

(vi) The value of land used for nonresidential purposes by the government. We compute this as $\mathcal{R}$ times the Current Cost Net Stock of Government Fixed Assets (line 8 of BEA Fixed Asset Table 1.1) less the current cost of infrastructure capital owned by the federal, state, and local government, as defined in Section B.2.2. We set the annual as the average of the current and lagged year-end observations.

We set $\sum(P_{bt}k_{bit} + p_{lit}l_{bit})$ equal to the aggregate value of all capital and finished land used in production, computed as the sum of the following variables:

(a) The total market value of tangible assets owned by nonfarm nonfinancial corporate businesses less the replacement cost of residential structures owned by nonfarm nonfinancial businesses, line 2 less line 33 of FFA Table B.102.[12] We set the annual as the average of the reported quarterly observations.

(b) The total market value of tangible assets less the market value of residential real estate owned by nonfarm nonfinancial noncorporate businesses,

[12] Residential structures are typically a very small fraction of total tangible assets: In 2009, they accounted for 1.4% of value.

line 2 less line 4 of FFA Table B.103. We set the annual as the average of the reported quarterly observations.

(c) The value of land used for nonresidential purposes by financial corporations computed in step (iv) above plus the Current Cost Net Stock of Private Nonresidential Fixed Assets owned by Financial Corporations, line 25 of BEA Fixed Asset Table 4.1. We set the annual as the average of the current and lagged year-end observations.

(d) The value of land used for nonresidential purposes by nonprofit organizations computed in step (v) above plus equipment and software owned by nonprofit organizations (line 6 of FFA Table B.100) plus the replacement cost of nonresidential structures owned by nonprofit organizations (line 46 of FFA Table B.100). We set the annual as the average of the reported quarterly observations.

(e) The value of land used for nonresidential purposes by the government computed in step (vi) above plus the Current Cost Net Stock of Government Fixed Assets (line 8 of BEA Fixed Asset Table 1.1) less the current cost of infrastructure capital owned by the federal, state, and local governments, as defined in Section B.2.2. We set the annual as the average of the current and lagged year-end observations.

(f) The Current Cost Net Stock of Consumer Durable Goods, line 13 of BEA Fixed Asset Table 1.1. We set the annual as the average of the current and lagged year-end observations.

We define $\mathcal{R}$ as the value of all land used for nonresidential purposes by businesses (the sum of items (i)–(iii) above) divided by the value of all tangible assets less land used for nonresidential purposes by businesses (the sum of (a) and (b) above less the sum of items (i)–(iii) above).

Also note that (as mentioned previously) when we use data from the BEA Fixed Asset Tables, we compute current-year values as the average of the reported current- and previous-year values. We do this because the BEA reports values at year end; this adjustment aligns the timing of the BEA data with that of the FFA data.

Data for Equation (S26). We compute

$$\frac{\sum w_{it}n_{it}}{\sum[w_{it}n_{it}+r_{lit}l_{bit}+r_{bt}k_{bit}]}$$

as follows. We set the numerator equal to "unambiguous labor income." We set the denominator equal to total gross domestic income plus an estimate of the nominal service flow from the stock of durable goods less the reported consumption of housing services less an estimate of "ambiguous income" (i.e., income that is neither unambiguous capital nor unambiguous labor income).

- We set unambiguous labor income equal to line 2 of Table 1.10 of the National Income and Product Accounts (NIPA), "Compensation of em-

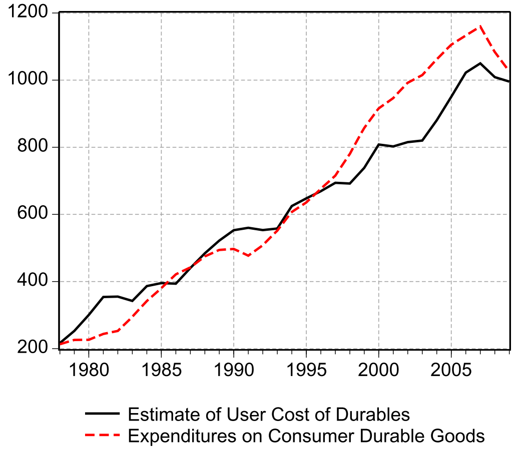

FIGURE S.1.—Comparison of user cost and expenditures on durable goods.

ployees, paid." This table is available at http://www.bea.gov/national/nipaweb/SelectTable.asp?Selected=N.

• We set gross domestic income equal to line 1 of NIPA Table 1.10, "Gross domestic income."

• We estimate the nominal service flow from the stock of durable goods as the sum of nominal depreciation on the stock of durable goods, line 13 of BEA Fixed Asset Table 1.3, plus the rate of interest on a 5-year Treasury bond times the nominal stock of durable goods. We take the nominal rate of interest on a 5-year Treasury bond from the website of the Federal Reserve Board, http://www.federalreserve.gov/releases/h15/data.htm. We set the nominal stock of durable goods as the average of the current- and previous- year reported (year-end) values of the stock, as reported in line 13 of BEA Fixed Asset Table 1.1.

Figure S.1 compares our estimate of the user cost of durables to expenditures on consumer durable goods as reported in the NIPA over our sample period, 1978–2009. Both data series are in billions of dollars. Broadly speaking, the levels and growth rates of the two series are similar.

• We set consumption of housing services equal to line 50 of NIPA Table 2.4.5, "Household consumption expenditures (for services): Housing."

• As in Cooley and Prescott (1995), we use the following data from NIPA Table 1.10 to determine ambiguous income:

line 9, Taxes on production and imports
− line 10, Subsidies
+ line 15, Proprietors' income with inventory valuation and CCA (capital consumption allowance)
+ line 22, Current surplus of government enterprises
+ line 26, Statistical discrepancy.

B.2.8. *Data Used for the Infrastructure Share of Finished Land*

One of our moment conditions for infrastructure's share of finished land rents, ζ, the growth rates of the price of finished land and infrastructure capital, g_{p_l} and g_{p_f}, and the depreciation rate on infrastructure capital, κ_f, is

$$E\left(\frac{R/g_{p_l}-(1-\kappa_f)^{\zeta}}{R/g_{p_f}-(1-\kappa_f)}\zeta-\frac{\sum P_{ft}k_{fit}}{\sum(p_{lit}l_{bit}+p_{lit}l_{hit})}\right)=0.$$

We set $\sum p_{ft}k_{fit}$ equal to the aggregate value of infrastructure capital, measured as defined in Section B.2.2. We set $\sum(p_{lit}l_{bit}+p_{lit}l_{hit})$ as the sum of the aggregate value of finished land used for business purposes, $\sum p_{lit}l_{bit}$, measured as defined in Section B.2.7, data for equation (S26), and the aggregate value of finished land used in housing, $\sum p_{lit}l_{hit}$, measured as defined in Section B.2.6.

B.2.9. *Data for the Growth of Aggregate per Capita Real Consumption*

Our moment condition for growth in aggregate real per capita consumption, g_c, is

$$E\{(\ln C_t-\ln(g_c)t)t\}=0.$$

We compute C_t as nominal aggregate consumption divided by the appropriate price index and divided again by the population.

We define the population as the civilian noninstitutional population aged 16 and older. These data are available from the Bureau of Labor Statistics. We replace the reported population with a predicted value based on a regression of the BLS data on a fourth order polynomial in year (R^2 of 0.998). This smoothes a few odd peaks in the reported BLS series.

We define nominal aggregate consumption as follows:

- total consumption as reported by the NIPA, line 1 of NIPA Table 2.4.5,
- less expenditures on durable consumption goods, line 3 of NIPA Table 2.4.5,
- less the consumption of housing services, line 50 of NIPA Table 2.4.5,
- plus government consumption expenditures, line 3 of NIPA Table 3.9.5,
- plus an estimate of the nominal service flow from the stock of durable goods. This estimate is described in detail in the previous section.

We compute the price index for this definition of consumption by chain-weighting the appropriate price indexes. For total consumption, expenditures on durable goods, and the consumption of housing services, the price indexes are available in NIPA Table 2.4.4. The price index for government consumption expenditures is available in NIPA Table 3.9.4. Finally, we set the price index for the service flow from durable goods equal to an estimate of the price index for the stock of durable goods. We estimate this as the average of the current and previous year values of the price index, which is computed as the reported nominal year-end stock of consumer durable goods (line 13 of BEA Fixed Asset Table 1.1) divided by the reported quantity index for this stock (line 13 of BEA Fixed Asset Table 1.2).

B.3. *Other Data*

In addition to the lagged endogenous panel variables in the system, we use two other MSA-level variables as instruments in our GMM analysis: per capita personal income, as measured by the BEA in Table CA1-3 of its Local Area Personal Income and Employment Tables, and repeat-sales price indexes for existing homes as produced by the Federal Housing Finance Agency. In our GMM analysis, we log and demean both variables. For the purposes of comparing simulated model output on employment and average wages across all 366 MSAs in the United States, we use data from the Local Area Personal Income and Employment Tables of the BEA. By MSA, wage and salary employment is reported on line 7020 of Table AMSA04 (Personal income and its components) and we compute average wage as the sum of "Wage and salary disbursements" (line 50) and "Supplements to wages and salaries" (line 60) divided by wage and salary employment. For the 22 MSAs in our sample, we compare the BEA-based average wage and total employment measures to estimates from the CPS, generated as total hours worked for all respondents (for employment) and average wage per hour for all hours worked by all workers. After removing year effects and taking logs, the correlation of the average wage estimates is 0.76 and for the employment estimates is 0.99.

APPENDIX C: DERIVATION OF AGGREGATE MOMENT CONDITIONS

In the main text, we assumed variables were demeaned when we stated moment conditions used to calculate trend rates of growth. In practice, we incorporate the estimates of the means in our calculations. The moment conditions stated here incorporate this estimation of the means.

We use the following moment conditions to identify κ_b, κ_s, and κ_f:

$$E\left\{\kappa_b - \frac{D_{bt}}{K_{bt}}\right\} = 0,$$

$$E\left\{\kappa_s - \frac{D_{st}}{K_{st}}\right\} = 0,$$

$$E\left\{\kappa_f - \frac{D_{bt}}{K_{ft}}\right\} = 0,$$

where D_{Xt} is nominal depreciation of capital of type $X = B, S, F$.

Along the balanced growth path (without aggregate uncertainty, but *with* idiosyncratic uncertainty) the household's Euler equation for *finished land* holds for each city,

$$p_{lit} = E_{t|i}\left\{\frac{1}{R}\left[r_{lit+1} + (1-\kappa_f)^{\zeta} p_{lit+1}\right]\right\}, \tag{S27}$$

where $E_{t|i}$ denotes expectation at t conditional on i, r_{li} is the rental price of finished land in city i, and p_{li} is the capital price of finished land in city i. This equation is derived as follows. Finished land k_{lit} is a Cobb–Douglas aggregate of raw land L_i (normalized to 1 in the main text) and infrastructure capital k_{fit}, such that

$$k_{lit} = l_i^{1-\zeta} k_{fit}^{\zeta}.$$

Raw land does not depreciate, but infrastructure capital depreciates at rate κ_f such that in the absence of any investment,

$$\frac{k_{fit+1}}{k_{fit}} = 1 - \kappa_f.$$

This implies that in the absence of any investment in infrastructure capital, finished land essentially depreciates. To see this, write

$$\begin{aligned} \frac{k_{lit+1}}{k_{lit}} &= \left(\frac{k_{fit+1}}{k_{fit}}\right)^{\zeta} \\ &= (1-\kappa_f)^{\zeta}. \end{aligned}$$

Consider the household raising its holdings of finished land in city i by k_{lit} units this period, and next period you rent the land and then resell it after it depreciates. The no arbitrage condition for this transaction is equation (S27).

Equation (S27) implies that along a balanced growth path, the average rental price of land and the average price of land grow at the same rate, g_{p_l}. We do not measure rents from land, but we do measure rents from housing, which includes land and structures. We assume housing services are derived from structures and finished land as

$$h_{it} = k_{sit}^{\omega} l_{hit}^{1-\omega}.$$

From profit maximization of housing service providers, at each date and in each city,

$$r_{hit} = \omega^{-\omega}(1-\omega)^{\omega-1} r_{lit}^{1-\omega} r_{st}^{\omega}, \tag{S28}$$

where r_h denotes the rent on services from houses and r_s denotes the rent on housing structures. Therefore, along a balanced growth path

$$r_{hit} = g_{p_l}^{(1-\omega)t} g_{p_s}^{\omega t} r_{hi0},$$
$$\ln E_t r_{hit} = \ln r_{hi0} + \big[(1-\omega)\ln(g_{p_l}) + \omega \ln(g_{p_s})\big]t,$$

since rents on housing structures and land follow the same trends as their respective asset prices.

Along a balanced growth path, for $x = b, s, f, l$,

$$P_{xt} = P_{x0} g_{p_x}^t,$$
$$\ln P_{xt} = \ln P_{x0} + \ln(g_{p_x})t,$$

where P_{xt} is the real price of the indicated type of capital (in the case of land, this is the average price). We identify g_{P_B}, g_{P_S}, g_{P_F}, and g_{P_L} using the moment conditions

$$E\big\{\ln P_{xt} - \ln P_{x0} - \ln(g_{p_x})t\big\} = 0, \quad x = b, s, f,$$
$$E\big\{\big(\ln P_{bt} - \ln P_{x0} - \ln(g_{p_x})t\big) \cdot t\big\} = 0, \quad x = b, s, f,$$
$$E\big\{\ln E_t r_{hit} - \ln E r_{hi0} - \big[(1-\omega)\ln(g_{p_l}) + \omega \ln(g_{p_s})\big]t\big\} = 0,$$
$$E\big\{\big(\ln E_t r_{hit} - \ln E_t r_{hi0} - \big[(1-\omega)\ln(g_{p_l}) + \omega \ln(g_{p_s})\big]t\big) \cdot t\big\} = 0.$$

The moment conditions for identifying g_{p_l} (the final two conditions) are based on the fact that $r_{lit} = g_{p_l}^t r_{li0}$ implies $E_t r_{lit} = g_{p_l}^t E_0 r_{li0}$, where E_t denotes expectation at t over i.

We now use equation (S27) evaluated along the balanced growth path to relate prices of land to rent from land. Analogous relationships hold for the other forms of capital. We use these relationships to formulate moment conditions to identify ω, α, ϕ, and ζ. Notice that along a balanced growth path,

$$E_t p_{lit+1} = g_{p_l} E_t p_{lit},$$
$$E_t r_{lit+1} = g_{p_l} E_t r_{lit}.$$

Since $E_t E_{t|i} x_t = E_t x_t$, it follows from (S27) that

$$E_t p_{lit} = \frac{g_{p_l}}{R - (1-\kappa_f)^\zeta g_{p_l}} E_t r_{lit}.$$

Similar conditions hold for the other types of capital.

We identify ζ, the share of development capital in the production of finished land, as follows. In every city,

$$\zeta = \frac{r_{fit} k_{fit}}{r_{lit} l_{bit} + r_{lit} l_{hit}},$$

so that

$$\zeta = \frac{E_t r_{fit} k_{fit}}{E_t r_{lit} l_{bit} + E_t r_{lit} l_{hit}}.$$

Using the relationship between values and incomes,

$$\frac{E_t P_{ft} k_{fit}}{E_t p_{lit} l_{bit} + E_t p_{lit} l_{hit}} = \frac{R/g_{p_l} - (1-\kappa_f)^{\zeta}}{R/g_{p_f} - (1-\kappa_f)} \frac{E_t r_{ft} k_{fit}}{[E_t r_{lit} l_{bit} + E_t r_{lit} l_{hit}]}.$$

Therefore, we identify ζ using

$$E\left\{\frac{R/g_{p_l} - (1-\kappa_f)^{\zeta}}{R/g_{p_f} - (1-\kappa_f)}\zeta - \frac{\sum P_{ft} k_{fit}}{\sum (p_{lit} l_{bit} + p_{lit} l_{hit})}\right\} = 0.$$

To identify ω, first relate the ratio of capital to land income in the housing sector:

$$\frac{r_{st} k_{sit}}{r_{lit} l_{hit}} = \frac{\omega}{1-\omega}.$$

Use this and the relationship between value and income ratios to obtain a relationship between the share of land in house values and ω:

$$\begin{aligned}\frac{E_t p_{lit} l_{hit}}{E_t p_{st} k_{sit} + E_t p_{lit} l_{hit}} &= \frac{1}{\dfrac{E_t P_{st} k_{sit}}{E_t p_{lit} l_{hit}} + 1} \\ &= \frac{1}{\dfrac{\omega}{1-\omega}\dfrac{R/g_{p_l} - (1-\kappa_f)^{\zeta}}{R/g_{p_s} + \kappa_s - 1} + 1}.\end{aligned}$$

The moment condition that identifies ω is then

$$E\left\{\frac{\sum p_{lit} l_{hit}}{\sum (P_{st} k_{sit} + p_{lit} l_{hit})}\left[\frac{\omega}{1-\omega}\frac{R/g_{p_l} - (1-\kappa_f)^{\zeta}}{R/g_{p_s} + \kappa_s - 1} + 1\right] - 1\right\} = 0.$$

Now consider the identification of α and ϕ. For this, we make use of the following relationships implied by the intermediate good producer's production function:

$$\phi(1-\alpha) = \frac{E_t w_{it} n_{it}}{E_t w_{it} n_{it} + E_t r_{lit} l_{bit} + E_t r_{bt} k_{bit}}, \tag{S29}$$

$$\frac{\alpha\phi}{1-\phi} = \frac{E_t r_{bt} k_{bit}}{E_t r_{lit} l_{bit}}.$$

We use the last equality to relate the ratio of the value of land to the value of tangible assets in the business sector to α and ϕ:

$$\frac{E_t p_{lit} l_{bit}}{E_t P_{bt} k_{bit} + E_t p_{lit} l_{bit}} = \frac{1}{\dfrac{E_t P_{bt} k_{bit}}{E_t p_{lit} l_{bit}} + 1} = \frac{1}{\dfrac{\alpha\phi}{1-\phi} \dfrac{R/g_{p_l} - 1}{R/g_{p_b} + \kappa_b - 1} + 1}.$$

Using the last equality and (S29), we arrive at the moment conditions used to identify α and ϕ:

$$E\left\{\frac{\sum p_{lit} l_{bit}}{\sum [P_{bt} k_{bit} + p_{lit} l_{bit}]}\left[\frac{\alpha\phi}{1-\phi} \frac{R/g_{p_l} - 1}{R/g_{p_b} + \kappa_b - 1} + 1\right] - 1\right\} = 0,$$

$$E\left\{\frac{\sum w_{it} n_{it}}{\sum [w_{it} n_{it} + r_{lit} l_{bit} + r_{bt} k_{bit}]} - \phi(1-\alpha)\right\} = 0.$$

We also need to estimate g_c, the gross growth rate of per capita consumption. Along a balanced growth path,

$$C_t = C_0 g_c^t,$$
$$\ln C_t = \ln C_0 + \ln(g_c) t,$$

where C_t is per capita consumption. Therefore, we identify g_c using the two moment conditions

$$E\{\ln C_t - \ln C_0 - \ln(g_c) t\} = 0,$$
$$E\{(\ln C_t - \ln C_0 - \ln(g_c) t) \cdot t\} = 0.$$

APPENDIX D: MEASURING THE EFFECT OF AGGLOMERATION ON PER CAPITA CONSUMPTION GROWTH

The balanced growth path of our model has

$$g_c = \gamma_a^{(1-\alpha)\delta/((1-\alpha)\delta+(\zeta-1)(\delta-1))} \gamma_n^{(1-\zeta)(\delta-1)/((1-\alpha)\delta+(\zeta-1)(\delta-1))} \times \gamma_b^{\alpha\delta/((1-\alpha)\delta+(\zeta-1)(\delta-1))} \gamma_f^{\zeta(1-\delta)/((1-\alpha)\delta+(\zeta-1)(\delta-1))}.$$

Use the balanced growth equation to express γ_a as

$$\hat{\gamma}_a = \hat{g}_c^{((1-\alpha)\delta+(\zeta-1)(\delta-1))/((1-\alpha)\delta)} \hat{\gamma}_n^{(1-\zeta)(1-\delta)/((1-\alpha)\delta)} \hat{g}_{p_b}^{\alpha/(1-\alpha)} \hat{g}_{p_f}^{\zeta(1-\delta)/((1-\alpha)\delta)},$$

$$g_c^* = \hat{\gamma}_a^{(1-\alpha)\phi/((1-\alpha)\phi+(\zeta-1)(\phi-1))} \gamma_n^{(1-\zeta)(\phi-1)/((1-\alpha)\phi+(\zeta-1)(\phi-1))} \times g_{p_b}^{-\alpha\phi/((1-\alpha)\phi+(\zeta-1)(\phi-1))} g_{p_f}^{-\zeta(1-\phi)/((1-\alpha)\phi+(\zeta-1)(\phi-1))}$$

$$
\begin{aligned}
&= \big[\hat{g}_c^{((1-\alpha)\delta+(\zeta-1)(\delta-1))/((1-\alpha)\delta)}\hat{\gamma}_m^{(1-\zeta)(1-\delta)/((1-\alpha)\delta)} \\
&\quad \times \hat{g}_{p_b}^{\alpha/(1-\alpha)}\hat{g}_{p_f}^{\zeta(1-\delta)/((1-\alpha)\delta)}\big]^{(1-\alpha)\phi/((1-\alpha)\phi+(\zeta-1)(\phi-1))} \\
&\quad \times \gamma_n^{(1-\zeta)(\phi-1)/((1-\alpha)\phi+(\zeta-1)(\phi-1))} g_{p_b}^{-\alpha\phi/((1-\alpha)\phi+(\zeta-1)(\phi-1))} \\
&\quad \times g_{p_f}^{-\zeta(1-\phi)/((1-\alpha)\phi+(\zeta-1)(\phi-1))} \\
&= \hat{g}_c^{\phi[(1-\alpha)\delta+(\zeta-1)(\delta-1)]/(\delta[(1-\alpha)\phi+(\zeta-1)(\phi-1)])}\hat{\gamma}_n^{(\phi-\delta)(1-\zeta)/(\delta[(1-\alpha)\phi+(1-\zeta)(1-\phi)])} \\
&\quad \times \hat{g}_{p_f}^{\zeta(\phi-\delta)/(\delta[(1-\alpha)\phi+(\zeta-1)(\phi-1)])}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\Lambda &= \frac{\hat{g}_c - g_c^*}{g_c^* - 1} \\
&= \big(\hat{g}_c - \hat{g}_c^{\phi[(1-\alpha)\delta+(\zeta-1)(\delta-1)]/(\delta[(1-\alpha)\phi+(\zeta-1)(\phi-1)])} \\
&\quad \times \hat{\gamma}_n^{(\phi-\delta)(1-\zeta)/(\delta[(1-\alpha)\phi+(1-\zeta)(1-\phi)])}\hat{g}_{p_f}^{\zeta(\phi-\delta)/(\delta[(1-\alpha)\phi+(\zeta-1)(\phi-1)])}\big) \\
&\quad \big/\big(\hat{g}_c^{\phi[(1-\alpha)\delta+(\zeta-1)(\delta-1)]/(\delta[(1-\alpha)\phi+(\zeta-1)(\phi-1)])} \\
&\quad \times \hat{\gamma}_N^{(\phi-\delta)(1-\zeta)/(\delta[(1-\alpha)\phi+(1-\zeta)(1-\phi)])}\hat{g}_{p_f}^{\zeta(\phi-\delta)/(\delta[(1-\alpha)\phi+(\zeta-1)(\phi-1)])} - 1\big).
\end{aligned}
$$

APPENDIX E: SOLVING THE MODEL AND COMPARING IT TO DATA

This section describes how we solve our model with $\omega = 0$ and $\xi = 1$. The notation is somewhat different from the main text of the paper, but is internally consistent.

In this model, there is a representative household with a large number of members who share consumption risk perfectly. Each period the household allocates its workers and capital across locations, chooses how much infrastructure capital to build in each location for the next period, and decides how much business capital it wants to allocate across locations in the next period. We assume these decisions are made after the household observes total factor productivity in each location in that period. The household takes all prices and the distribution of total factor productivity as given. That is, it behaves competitively and does not take into account the effect of its actions on the density of production in each location. In each location there are developers and producers. The developers rent local infrastructure capital from the household and combine it with raw land to produce developed land, which they then rent to local producers. Producers rent capital and labor from the household and developed land to produce the city-specific intermediate good. There is also a final good producer who combines city-specific intermediate goods to produce the final good. Developers and goods producers maximize profits, taking all prices and total factor productivity as given.

The competitive equilibrium for this economy can be found as the solution to an optimization problem with side conditions. Notation is borrowed from the main text except that here we assume that the support of the distribution of technology is discrete. Let $m(z^t)$ denote the distribution of cities across idiosyncratic productivity histories z^t. The household perfectly insures itself against consumption risk, so we write the optimization problem as

$$\max_{\{C_t,K_{t+1},y(z^t),l(z^t),k(z^t),n(z^t),h(z^t),x(z^t),d(z^t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \sum m(z^t)n(z^t) \times \left[\ln C_t + \psi \ln h(z^t)\right]$$

subject to

$$C_t + K_{t+1} - (1-\kappa)K_t + X_t \leq \left[\sum m(z^t)y(z^t)^\eta\right]^{1/\eta},$$

$$y(z^t) \leq z_t^{(1-\alpha)\phi}\left[\frac{\hat{y}(z^t)}{\hat{l}(z^t)}\right]^{(\lambda-1)/\lambda} l(z^t)^{1-\phi} k(z^t)^{\alpha\phi} n(z^t)^{(1-\alpha)\phi},$$

$$\sum m(z^t)k(z^t) \leq K_t,$$

$$n(z^t)h(z^t) + l(z^t) \leq d(z^{t-1})^s, \quad \forall z^t,$$

$$\sum m(z^t)n(z^t) \leq 1,$$

$$d(z^t) \leq (1-\zeta)d(s^{t-1}) + x(z^t), \quad \forall z^t,$$

$$\sum m(z^t)x(z^t) \leq X_t,$$

$$K_0, d(z_0)\ \forall z_0, \hat{y}(z^t) \text{ and } \hat{l}(z^t) \text{ given.}$$

The un-indexed summations are over productivity histories z^t. The competitive equilibrium corresponds to a solution to this optimization problem such that $y(z^t) = \hat{y}(z^t)$ and $l(z^t) = \hat{l}(z^t)$. Note that for simplicity, and in contrast to the main text, we have assumed that housing services are derived from developed land only.

E.1. *Steady State With One City*

This section derives the steady state with one city, which we use for starting values in the general case. The first order condition and constraints for land development, evaluated in steady state, are

$$1 = \frac{1}{R}\left[r_d + (1-\zeta)\right],$$

$$x = \zeta d,$$

where r_d is the rental rate for infrastructure capital. Then

$$1 = \frac{1}{R}\big[r_l \varsigma d^{\varsigma-1} + (1-\zeta)\big],$$
$$r_d^* = R - 1 + \zeta,$$
$$R = 1/\beta,$$
$$r_l \varsigma d^{\varsigma-1} = r_d^*,$$

where r_l is rent on finished land.

Output in the city is

$$y = z^{(1-\alpha)\phi}\left[\frac{y}{l}\right]^{(\lambda-1)/\lambda} l^{1-\phi} k^{\alpha\phi} n^{(1-\alpha)\phi}.$$

Let $\delta = \lambda\phi$. Solving for output,

$$y = z^{\delta(1-\alpha)} l^{1-\delta} k^{\delta\alpha} n^{\delta(1-\alpha)}. \tag{S30}$$

Aggregate output is

$$Y = y$$

and output prices are

$$q = 1.$$

The labor supply constraint implies

$$n = 1$$

and the aggregate resource constraint is

$$C + \kappa K + \zeta D = Y,$$

where the total stock of developed land is equal to the developed land in the city, $D = d$. The equilibrium rent on business capital r_k^* is found using the business capital accumulation and is given by

$$r_k^* = 1/\beta - 1 + \kappa.$$

Rental demand for capital, developed land demand, labor demand, housing, land development, labor allocation first order necessary conditions (FONCS), and the land constraint are

$$k = \alpha\phi y r_k^{*-1},$$
$$r_l = y(1-\phi)l^{-1},$$

$$w = (1-\alpha)\phi y n^{-1},$$
$$\psi C = r_l h,$$
$$r_l \varsigma d^{\varsigma-1} = r_d^*,$$
$$\psi \ln h = \frac{1}{C}\theta + \frac{1}{C}[\psi C - w],$$
$$nh + l = d^{\varsigma}.$$

Here w denotes the wage. From the infrastructure FONC,

$$d^{\varsigma} = \frac{r_d^* d}{\varsigma r_l}.$$

Since $n = 1$, the land constraint can be written

$$l = \frac{r_d^* d}{\varsigma r_l} - h.$$

Using this, the land demand FONC and the housing demand FONC, we have

$$y(1-\phi) = \frac{r_d^* d}{\varsigma} - \psi C. \tag{S31}$$

Substitute from the capital rental and land demand FONCS into (S30) to express y as function of r_l:

$$\begin{aligned} y &= z^{\delta(1-\alpha)}\left[\frac{(1-\phi)y}{r_l}\right]^{1-\delta}\left[\frac{\alpha\phi}{r_k^*}y\right]^{\delta\alpha} \\ &= \left[z^{\delta(1-\alpha)}[1-\phi]^{1-\delta}\left[\frac{\alpha\phi}{r_k^*}\right]^{\delta\alpha}\right]^{1/(\delta(1-\alpha))} r_l. \end{aligned}$$

Use the definition of r_d to solve for d as a function of r_l:

$$d = \left[\frac{\varsigma}{r_d^*}\right]^{1/(1-\varsigma)} r_l.$$

Solve for k as a function of r_l using the capital rental FONC:

$$k = \frac{\alpha\phi}{r_k^*}\left[z^{\delta(1-\alpha)}[1-\phi]^{1-\delta}\left[\frac{\alpha\phi}{r_k^*}\right]^{\delta\alpha}\right]^{1/(\delta(1-\alpha))} r_l.$$

Solve for C as a function of y and d using (S31):

$$C = \frac{r_d^*}{\psi\varsigma}d - \frac{1-\phi}{\psi}y.$$

With this last expression, we solve the aggregate resource constraint for r_l using the expressions for y and d as functions of r_l from above. First substitute in for the right hand side of the aggregate resource constraint:

$$\begin{aligned} y &= C + \kappa k + \zeta d \\ &= \frac{r_d^*}{\psi\varsigma} d - \frac{1-\phi}{\psi} y + \kappa \frac{\alpha\phi}{r_k^*} y + \zeta d. \end{aligned}$$

Substituting for y and d in the last equation and solving for r_l, we arrive at

$$\begin{aligned} r_l^* = \Bigg[&\left[\frac{r_d^*}{\varsigma\psi} + \zeta \right] \left[\frac{\varsigma}{r_d^*} \right]^{1/(1-\varsigma)} \\ &\Bigg/ \left(\left[1 + \frac{1-\phi}{\psi} - \kappa \frac{\alpha\phi}{r_k^*} \right] \right. \\ &\times \left. \left[z^{\delta(1-\alpha)} [1-\phi]^{1-\delta} \left[\frac{\alpha\phi}{r_k^*} \right]^{\delta\alpha} \right]^{1/(\delta(1-\alpha))} \right) \Bigg]^{1/((\delta-1)/(\delta(1-\alpha))-1/(1-\varsigma))}. \end{aligned}$$

With r_l^* in hand, we can solve for C^*, k^*, d^*, and y^* using expressions derived above. In addition, from the housing FONC,

$$h^* = \psi C^* / r_l^*,$$

and from the land constraint,

$$l^* = r_d^* d^* / r_l^* - h^*.$$

Finally, the labor demand and labor allocation FONCS yield

$$\begin{aligned} w^* &= y^*(1-\alpha)\phi, \\ \theta^* &= C^*\psi \ln h^* - \psi C^* + w^*. \end{aligned}$$

The foregoing demonstrates that, for admissible parameters, a nonstochastic steady state exists and is unique with one level of productivity.

E.2. *General Steady State Solution*

The solution strategy is to fix C, K, D, and θ, solve for all other endogenous variables conditional on these variables, and then find a fixed point in C, K, D, and θ that satisfies market clearing.

Using the transition equation for exogenous productivity it is straightforward to compute the steady state distribution of cities by exogenous productivity. The mass of cities of type i is m_i and the aggregate output is

$$Y = \left[\sum m_i y_i{}^{\eta} \right]^{1/\eta}. \tag{S32}$$

Output prices are

$$q_i = Y^{1-\eta} y_i^{\eta-1}.$$

The aggregate resource constraint is

$$C + \kappa K + \zeta D = Y,$$

so that

$$q_i = [C + \kappa K + \zeta D]^{1-\eta} y_i^{\eta-1}.$$

Except where it helps the exposition, we drop the i subscript hereafter. The FONC for infrastructure capital accumulation is

$$1 = \beta E\big[r_l \varsigma d'^{\varsigma-1} + 1 - \zeta\big].$$

From the housing FONC and land constraint,

$$r_l = \frac{\psi C n}{d^{\varsigma} - l}.$$

Therefore, the FONC for infrastructure capital can be written

$$1 = \beta E\left[\frac{\psi C n'}{d'^{\varsigma} - l'} \varsigma d'^{\varsigma-1} + 1 - \zeta\right].$$

Equilibrium r_k is given by

$$r_k^* = 1/\beta - 1 + \kappa.$$

From the FONCS and constraints,

(S33) $$n_i h_i + l_i = d_i^{\varsigma},$$

(S34) $$\psi C = r_l h_i,$$

(S35) $$\psi C \ln h_i = \theta + [\psi C - w_i],$$

(S36) $$w_i = q_i(1-\alpha)\phi y_i n_i^{-1},$$

(S37) $$r_{li} = q_i(1-\phi) y_i l_i^{-1},$$

(S38) $$q_i = [C + \kappa K + \zeta D]^{1-\eta} y_i^{\eta-1},$$

(S39) $$r_k^* = q_i \alpha \phi y_i k_i^{-1},$$

(S40) $$y_i = z_i^{\delta(1-\alpha)} l_i^{1-\delta} k_i^{\delta\alpha} n_i^{\delta(1-\alpha)},$$

(S41) $$1 = \beta E\left[\frac{\psi C}{h_i'} \varsigma d_i'^{\varsigma-1} + 1 - \zeta\right],$$

(S42) $$k_i = q_i \alpha \phi y_i r_k^{*-1}.$$

Here we have used the i subscript to be clear on which variables are location specific and which are common to all locations. Combine all but (S38) and (S40), (S41), (S42) to form

$$\frac{\psi C}{r_k^* h} = \frac{(1-\phi)}{\alpha\phi}\frac{k}{l}, \tag{S43}$$

$$\frac{(1-\alpha)k}{\alpha n} = \frac{\theta + \psi C - \psi C \ln h}{r_k^*}, \tag{S44}$$

$$nh + l = d^{\varsigma}. \tag{S45}$$

It follows that

$$\frac{l\psi C}{r_k^*(d^{\varsigma} - l)} = \frac{(1-\phi)}{\alpha\phi}\frac{k}{n}. \tag{S46}$$

Using (S43), we can rewrite the infrastructure FONC as

$$1 = \beta E\left[\frac{(1-\phi)r_k^*}{\alpha\phi}\frac{k}{l}\varsigma d'^{\varsigma-1} + 1\right].$$

Furthermore, combining (S44) and (S46),

$$\frac{l\psi C}{r_k^*(d^{\varsigma} - l)} = \frac{(1-\phi)}{(1-\alpha)\phi}\frac{\theta + \psi C - \psi C \ln h}{r_k^*}.$$

Solving this last equation for h yields

$$\begin{aligned} h &= \exp\left(\frac{\theta}{\psi C} + 1 - \frac{(1-\alpha)\phi l}{(d^{\varsigma} - l)(1-\phi)}\right) \\ &= \hat{h}(d, l; C, \theta). \end{aligned}$$

Using (S43),

$$\begin{aligned} k &= \frac{\alpha\phi\psi C}{(1-\phi)r_k^*\hat{h}(d, l; C, \theta)} l \\ &= \hat{k}(d, l; C, \theta). \end{aligned}$$

Using (S33),

$$\begin{aligned} n &= \frac{d^{\varsigma} - l}{\hat{h}(d, l; C, \theta)} \\ &= \hat{n}(d, l; C, \theta). \end{aligned}$$

Using (S38), (S39), and (S40), and substituting in the expressions for k and n, we solve for l_j for each d'_i using

$$r_k^* = \alpha\phi[C + \kappa K + \zeta D]^{1-\eta} z_j^{\delta(1-\alpha)\eta} l_j^{(1-\delta)\eta} \times \hat{k}(d'_i, l_j; C, \theta)^{\delta\alpha\eta-1} \hat{n}(d'_i, l_j; C, \theta)^{\delta(1-\alpha)\eta}.$$

The "prime" superscript denotes the choice of the indicated variable made for the following period when the current state is given by the subscript. For each i, this yields

$$l_{ji} = \tilde{l}(C, K, D, \theta, z_j, d'_i), \tag{S47}$$

$$n_{ji} = \tilde{n}(C, K, D, \theta, z_j, d'_i), \tag{S48}$$

$$k_{ji} = \tilde{k}(C, K, D, \theta, z_j, d'_i). \tag{S49}$$

Here we use the subscript convention that the variable is chosen contemporaneously with the technology state that corresponds to the first subscript and the state that corresponds to the second subscript in the period before.

From the land infrastructure FONC, for each i,

$$\begin{aligned} 1 &= \beta \sum \pi_{ij} \left[\frac{(1-\phi) r_k^*}{\alpha\phi} \frac{\tilde{k}(C, K, D, \theta, z_j, d'_i)}{\tilde{l}(C, K, D, \theta, z_j, d'_i)} \varsigma d_i'^{\varsigma-1} + 1 - \zeta \right] \\ &= \beta \sum \pi_{ij} \big[f(C, K, D, \theta, z_j, d'_i) + 1 - \zeta \big]. \end{aligned} \tag{S50}$$

Here the π_{ij} denotes the probability of transitioning from state i to state j. Because π_{ij} depends on z_i, solving (S50) yields

$$d'_i = d(C, K, D, \theta, z_i).$$

Using (S47), (S48), and (S49), we may now obtain

$$l_{ij} = l(C, K, D, \theta, z_i, z_j),$$

$$n_{ij} = n(C, K, D, \theta, z_i, z_j),$$

$$k_{ij} = k(C, K, D, \theta, z_i, z_j).$$

The switch in the order of the subscripts is by design. These expressions can be substituted into (S40) to construct

$$y_{ij} = y(C, K, \theta, D, z_i, z_j).$$

By the definition of developed land investment,

$$\begin{aligned} x_{ij} &= d'_i - (1-\zeta) d_j \\ &= x(C, K, D, \theta, z_i, z_j). \end{aligned}$$

In addition,

$$\begin{aligned} Y &= C + \kappa K + \zeta D \\ &= Y(C, K, X). \end{aligned}$$

We solve for C, K, D, and θ using

$$\begin{aligned} &Y(C, K, D, \theta) = \Big[\sum m_{ij} y(C, K, D, \theta, z_i, z_j)^\eta\Big]^{1/\eta}, \\ &K = \sum m_{ij} k(C, K, D, \theta, z_i, z_j), \\ &1 = \sum m_{ij} n(C, K, D, \theta, z_i, z_j), \\ &X = \sum m_{ij} x(C, K, D, \theta, z_i, z_j), \end{aligned}$$

where m_{ij} denotes the steady state mass of cities that have z_i today and z_j yesterday. Notice that the city level variables are entirely determined by lagged and current technology. In particular, for each possible pair of technology states there is a unique value of infrastructure capital chosen for the next period. Consequently, the number of infrastructure capital states corresponds to the number of possible pairs of technology states.

We use the GAUSS nonlinear equation solver `eqSolve` to solve this system of four equations. To evaluate these equations, we need to solve (S50). We accomplish this using a version of `eqSolve` that we have modified to exploit the sparseness of the transition matrix formed with the π_{ij} (see below).

For this, we need to find m_{ij}, the steady state distribution of (z_t, z_{t-1}). For simplicity, consider the case of three z states. The results described below are based on a grid for z_t with 75 points, so there are 5625 (z_t, z_{t-1}) states. We need the transition probabilities for (z_t, z_{t-1}) to (z_{t+1}, z_t). The state is summarized by

$$\begin{bmatrix} z_1 & z_1 \\ z_1 & z_2 \\ z_1 & z_3 \\ z_2 & z_1 \\ z_2 & z_2 \\ z_2 & z_3 \\ z_3 & z_1 \\ z_3 & z_2 \\ z_3 & z_3 \end{bmatrix}.$$

We obtain the π_{ij} (defined above) from the underlying transition matrix for z_t. Then the matrix of transition probabilities is

$$\begin{bmatrix} \pi_{11} & 0 & 0 & \pi_{12} & 0 & 0 & \pi_{13} & 0 & 0 \\ \pi_{11} & 0 & 0 & \pi_{12} & 0 & 0 & \pi_{13} & 0 & 0 \\ \pi_{11} & 0 & 0 & \pi_{12} & 0 & 0 & \pi_{13} & 0 & 0 \\ 0 & \pi_{21} & 0 & 0 & \pi_{22} & 0 & 0 & \pi_{23} & 0 \\ 0 & \pi_{21} & 0 & 0 & \pi_{22} & 0 & 0 & \pi_{23} & 0 \\ 0 & \pi_{21} & 0 & 0 & \pi_{22} & 0 & 0 & \pi_{23} & 0 \\ 0 & 0 & \pi_{31} & 0 & 0 & \pi_{32} & 0 & 0 & \pi_{33} \\ 0 & 0 & \pi_{31} & 0 & 0 & \pi_{32} & 0 & 0 & \pi_{33} \\ 0 & 0 & \pi_{31} & 0 & 0 & \pi_{32} & 0 & 0 & \pi_{33} \end{bmatrix}.$$

From this matrix, we can calculate the steady state m_{ij}. Obtaining the steady state and evaluating conditional expectations is greatly accelerated by exploiting the sparseness of this matrix.

E.3. *Approximating the Technology Process*

We consider an underlying technology process

$$\ln z_t = \max\{\gamma_z + \ln z_{t-1} + \varepsilon_t, \ln z_{\min}\}, \tag{S51}$$

where ε_t is i.i.d. normally distributed with mean zero and variance σ_ε^2. This is isomorphic to the process considered by Gabaix (2000). As long as $g < 0$ for fixed $z_{\min}$, this process has an invariant distribution in z_t, which is convenient for solving our model. The tail of this distribution is exponential, that is, it has the property that

$$\Pr[z_t > b] = \frac{a}{b^\vartheta}$$

for some $a > 0$ and $\vartheta > 0$. Zipf's law is $\vartheta = 1$ for the z_t population. We can always find a γ_z to match an admissible ϑ. The parameter ϑ can be estimated in our data by regressing the log rank z_t on $\ln z_t$. The coefficient on $\ln z_t$ is a consistent estimate of ϑ. Technology is well approximated by an exponential distribution with $\vartheta = 2.5$.

We approximate (S51) with a discrete Markov chain. The grid is chosen to be equally spaced in logs. The elements of the transition equation are conditional probabilities of transiting from a given grid point to intervals around all possible grid points, where the intervals are equally spaced from the midpoints between grid points. These probabilities are calculated using (S51). In practice, we assume wide domains for the grid, varying from 150 to over 300 times the standard deviation of the innovation. This ensures that mass does not accumulate on the largest grid points. We solve for the γ_z that yields $\vartheta = 20$ in the steady state, excluding a small number of the smallest and largest grid points. We exclude some grid points to focus on the region of the state space where the approximation is best. This strategy yields a remarkably good approximation to an exponential distribution, except in the extreme tails. In fact, the overall approximation resembles empirical plots of log city rank by population versus log population, with a different slope of course. We set $z_{\min} = 1 - 1/\zeta$. This sets the mean of the distribution to approximately 1. We assume $\sigma_\varepsilon = 0.001$.

APPENDIX F: STANDARD ERRORS

We estimate Λ in three steps. To begin, we collect the expressions in the moment conditions described in the last subsection into a vector-valued function $\Psi_1(X_t, \theta_1)$, so that

$$E\Psi_1(X_t, \theta_1) = 0. \tag{S52}$$

Here, X_t is a vector of the aggregate variables included in these moment conditions, and θ_1 is a parameter vector given by

$$\theta_1 \equiv [\kappa_b, \kappa_b, \kappa_b, g_{p_l}, g_{p_b}, g_{p_s}, g_{p_f}, \gamma_n, g_c, \alpha, \phi, \omega, \zeta]'.$$

Because this system of moment conditions is exactly identified, the dimensions of Ψ_1 and θ_1 are equal. The first step is to estimate equation (S34) by GMM, in which we use a Newey–West weight matrix with a lag length of 2.

In the second step, we estimate δ and ξ using the moment conditions in equation (26) in the main text. This estimation requires that we plug in the estimates of ω and α from the first step into (26). To account for the sampling variation associated with these two plug-in parameters, we adjust the weight matrix using the methods described in Newey and McFadden (1994). Specifically, write the moment condition in (26) as

$$E\Psi_2(X_{it}, \theta_2) = 0, \tag{S53}$$

in which X_{it} is the vector of panel data in (26) and $\theta_2 = \{\delta, \xi, \omega, \alpha\}$. Next, let $\psi_\omega(X_t)$ and $\psi_\alpha(X_t)$ be the influence functions associated with ω and α. To express the optimal weight matrix for the GMM estimation based on the moment condition in (S53), we define

$$\tilde{\Psi}_2(X_{it}, \theta_2) \equiv \Psi_2(X_{it}, \theta_2) - \sqrt{N}\left(\frac{\partial \Psi_2(X_{it}, \theta_2)}{\partial \omega}\psi_\omega(X_t) + \frac{\partial \Psi_2(X_{it}, \theta_2)}{\partial \alpha}\psi_\alpha(X_t)\right). \tag{S54}$$

The $\sqrt{N}$ term appears in this expression to account for the fact that the parameters ω and α are estimated with only T observations, instead of with NT observations. Finally, the optimal weight matrix is given by

$$\boldsymbol{\Omega} \equiv E\big[\tilde{\Psi}_2(X_{it}, \delta, \xi, \omega, \alpha)\tilde{\Psi}_2(X_{it}, \delta, \xi, \omega, \alpha)'\big].$$

The rest of the GMM estimation proceeds by averaging the moment conditions over both i and t, and by clustering the weight matrix at the city level.

The third step is to substitute the point estimates for g_c, g_{p_f}, δ, α, ϕ, and ζ into equation (17) in the main text to obtain Λ. To calculate the sampling variance of Λ, we need the joint covariance matrix of these six parameters, which we calculate by stacking the parameters' influence functions as shown by Erickson and Whited (2002). As in (S54), we multiply the influence functions for the parameters estimated with time-series data by $\sqrt{N}$. After this calculation, a standard application of the delta method gives the variance of Λ.

APPENDIX G: MONTE CARLO STUDY

We perform a Monte Carlo study of our estimator using simulated data whose distribution closely approximates that of our own data set, which consists of some variables that vary in only the time dimension and others that vary in both the time and the city dimension. We first consider the times-series variables

$$\text{(S55)}\qquad \mathbf{x}_t \equiv \begin{pmatrix} \dfrac{D_{bt}}{K_{bt}}, \dfrac{D_{st}}{K_{st}}, \dfrac{D_{ft}}{K_{ft}}, Er_{hit}, p_{bt}, p_{st}, p_{ft}, c_t, \\ \dfrac{\sum p_{lit} l_{hit}}{\sum (p_{st}k_{sit} + p_{lit}l_{hit})}, \dfrac{\sum p_{lit} l_{bit}}{\sum (p_{bt}k_{bit} + p_{lit}l_{bit})}, \\ \dfrac{\sum w_{it} n_{it}}{\sum (w_{it}n_{it} + r_{lit}l_{bit} + r_{bt}k_{bit})}, \dfrac{\sum p_{ft} k_{fit}}{\sum (p_{lit}l_{bit} + p_{lit}l_{hit})} \end{pmatrix}.$$

As a first step, we use our actual data to estimate a time-trend regression for $\mathbf{x}_t$,

$$\text{(S56)}\qquad \mathbf{x}_t = \mathbf{a} + \mathbf{b}t + \mathbf{u}_{xt},$$

in which $\mathbf{b}$ is a vector of time trends for the individual elements of $\mathbf{x}_t$, $\mathbf{a}$ is a vector of intercepts, and $\mathbf{u}_{xt}$ is a vector of disturbances with covariance matrix Σ_{ux}. With the estimates of $(\mathbf{a}, \mathbf{b}, \Sigma_{ux})$, we simulate each variable as follows. We generate a matrix of normal disturbances of length 132 and width equal to the dimension of $\mathbf{x}_t$. These disturbances are serially uncorrelated, but are contemporaneously correlated with a covariance matrix of Σ_u. We then generate $\mathbf{x}_t$ using (S56). Finally, we keep the last 32 observations, where 32 is the time span of our actual data set. This procedure give us time-series variables with the same first and second moments as those in our actual data.

Next we describe our panel variables $\Delta\hat{\mathbf{y}}_{it} = (\Delta\hat{w}_{eit}, \Delta\hat{r}_{hit}, \Delta\hat{p}_{yit}, \Delta\hat{s}_{it}, \Delta\hat{m}_{it}, \Delta\hat{v}_{it})$, in which $\Delta\hat{v}_{it}$ is a vector that contains our two additional instrumental variables, house prices, and per capita income. We simulate directly in first-differenced, "hatted" form. We first calculate the means and covariances of these variables in our actual data. We also calculate the first-order serial correlations from OLS estimates of a simple AR(1) model,

$$\Delta\hat{\mathbf{y}}_{it} = A\Delta\hat{\mathbf{y}}_{it-1} + \mathbf{u}_{yit}, \tag{S57}$$

in which A is a diagonal matrix of autoregressive coefficients. We denote the estimated covariance matrix of the residuals as Σ_y.

Next, we use these estimates to create simulated panel variables. First, we generate a matrix of normal disturbances, $\tilde{\mathbf{u}}_{yit}$, of length 132 and width equal to the dimension of $\Delta\hat{\mathbf{y}}_{it}$ times 22, which is the number of cities in our panel. We then update the variables $(\Delta\hat{r}_{hit}, \Delta\hat{p}_{yit}, \Delta\hat{s}_{it}, \Delta\hat{m}_{it}, \Delta\hat{v}_{it})$ in each of these cities using (S57). Finally, we construct $\Delta\hat{w}_{eit}$ using

$$\begin{aligned}\Delta\hat{w}_{eit} = {} & \frac{1}{1-\omega}\frac{\delta-1}{\delta(1-\alpha)}\Delta\hat{r}_{hit} + \frac{1}{\delta(1-\alpha)}\Delta\hat{q}_{it} \\ & + \frac{1-\xi}{\xi}\Delta\hat{\chi}_{it} + (\xi-1)\Delta\hat{m}_{it} + \varepsilon_{it},\end{aligned} \tag{S58}$$

in which ε_{it} is constructed as $\varepsilon^*_{it} + a'\tilde{\mathbf{u}}_{yit}$, in which a is a vector of coefficients that correspond to $(\Delta\hat{r}_{hit}, \Delta\hat{p}_{yit}, \Delta\hat{s}_{it}, \Delta\hat{m}_{it})$ and in which ε^*_{it} an i.i.d. normal variable. Thus, the error term in (S58) shares common contemporaneous variation with $(\Delta\hat{r}_{hit}, \Delta\hat{p}_{yit}, \Delta\hat{s}_{it}, \Delta\hat{m}_{it})$, as our model predicts. We set the variance of ε^*_{it} so that the variance of $\Delta\hat{w}_{eit}$ in our simulated data equals the variance of $\Delta\hat{w}_{eit}$ in our real data. We set the correlation parameters, a, so that the covariance between $\Delta\hat{w}_{eit}$ and $(\Delta\hat{r}_{hit}, \Delta\hat{p}_{yit}, \Delta\hat{s}_{it}, \Delta\hat{m}_{it})$ in our simulated data approximates this covariance in our actual data.

Finally, we note that all of the panel variables are the residuals from regressing the raw variables on time dummies. They are, therefore, by construction orthogonal to any time-series variables, so we set the covariances between the time-series and the panel variables equal to zero.

We repeat this procedure 10,000 times (thus generating 10,000 data sets), where we set the true values of the coefficients equal to our estimates from Table III. Specifically, $\delta = 1.04$, $\xi = 0.54$, and the time-series coefficients all equal their estimated values. We use these true values to evaluate $\Lambda = 0.102$.

We estimate the model using twice lagged values of $(\Delta\hat{w}_{eit}, \Delta\hat{r}_{hit}, \Delta\hat{q}_{it}, \Delta\hat{\chi}_{it}, \Delta\hat{m}_{it}, \Delta\hat{v}_{it})$. We average our GMM moment conditions in both the cross-sectional and time-series dimensions. We calculate standard errors as we do for Table I, with a Newey–West correction for the first stage time-series estimation and clustering at the city level for the second stage panel estimation.

TABLE S.XV

MONTE CARLO RESULTS

	δ	ξ	Λ
Coefficients			
Average coefficient	1.0414	0.5493	0.1021
Average bias	0.0014	0.0093	0.0013
Mean absolute deviation	0.0366	0.0319	0.0600
RMSE	0.0475	0.0505	0.0803
Test Statistics			
Upside null rejection frequency (2.5%)	0.0300	0.0468	0.0083
Downside null rejection frequency (2.5%)	0.0823	0.0542	0.0907
J-test 5% rejection rate	0.0102		
$m2$-test 5% rejection rate	0.0968		
$m3$-test 5% rejection rate	0.0000		

We report the results from this simulation in Table S.XV.

The top panel of Table S.XV reports the average estimated coefficient over the 10,000 trials, as well as the average bias, mean absolute deviation, and root mean square error (RMSE). We see that despite the small sample size, our two-step GMM estimator produces nearly unbiased coefficient estimates. The mean absolute deviations and RMSEs are low for δ and ξ, but somewhat larger for Λ. This result makes sense inasmuch as Λ is estimated from both time-series and panel data, and the time-series data contain much less variation identifying information.

The bottom panel of the table reports the two tail probabilities from nominal 5% tests that the coefficient estimates equal their true values. "Upside" refers to the right tail and "downside" refers to the left tail. In general, these tests are slightly oversized, which arises because the standard errors are "too small." However, the overrejection is not symmetric. For δ and Λ, the probability of rejecting the null on the upside is much smaller than the probability of rejecting the null on the downside. For δ, the right tail is approximately correctly sized, and for Λ, the right tail is undersized, which means that there is a negligible probability of rejecting the null in favor of the alternative of $\Lambda > 0$. Of course, the right tails are the ones that matter for our application because our theory implies that both δ and Λ are positive. Our test rejection results are, therefore, comforting in that they imply that our significant coefficients are not an artifact of a test overrejection in small samples.

Finally, we report the rejection rates for our three diagnostic tests. We find that the J-test and the $m3$-test underreject, but that the $m2$-test overrejects slightly. The first two results imply that the J-test and the $m3$-test are unlikely to be useful specification tests. In contrast, the third result implies that the insignificant $m2$-test statistic we find in our estimation is not likely to be an artifact of an undersized test.

APPENDIX H: EQUILIBRIUM EXISTENCE AND UNIQUENESS WITHOUT HOUSING

This section shows, for several versions of the model without housing, that there is a unique solution in which all cities are occupied when a simple parametric restriction is satisfied.

To simplify the notation, we index cities by their location on the unit interval rather than productivity history as we do in the paper and elsewhere in this document. Note that some of the notation is inconsistent with the main text and elsewhere in this document.

Allocations are obtained by solving an optimization problem, taking as given the distribution of TFP across locations i, A_i. The TFP equals a productivity shock multiplied by an increasing function of output per unit of land. An equilibrium is a solution to the optimization problem in which the implied TFP distribution equals the one assumed. Throughout, we assume that the exogenous productivity process z is a well behaved finite dimensional discrete Markov chain.

H.1. *Model Without Housing and Perfectly Substitutable Goods*

Equilibrium is the solution to the program

$$\max_{C,N_i,L_i}\Bigg\{\ln C + \pi\left[\int A_i N_i^{\alpha} L_i^{1-\alpha}\,di - C\right] + \pi\int \varphi_i[1-L_i]\,di + \pi\theta\left[1-\int N_i\,di\right]\Bigg\},\quad A_i \text{ given}$$

such that

$$A_i = \left[\frac{A_i N_i^{\alpha} L_i^{1-\alpha}}{L_i}\right]^{(\lambda-1)/\lambda}.$$

Note first of all that for given A_i, the program has a unique interior solution.

The necessary and sufficient conditions for an interior solution to the program for given A_i are

$$A_i \alpha N_i^{\alpha-1} L_i^{1-\alpha} = \theta,$$
$$1 = L_i,$$

$$A_i(1-\alpha)N_i^{\alpha}L_i^{-\alpha} = \varphi_i,$$

$$1 = \int N_i\, di,$$

$$C = \int A_i N_i^{\alpha} L_i^{1-\alpha},$$

$$\frac{1}{C} = \pi.$$

An equilibrium requires

$$A_i = z_i\big[A_i N_i^{\alpha}\big]^{1-1/\lambda} = z_i^{\lambda} N_i^{\alpha\lambda-\alpha}.$$

Substitute for A_i in the FONC for labor:

$$z_i^{\lambda} N_i^{\alpha\lambda-\alpha} \alpha N_i^{\alpha-1} = \theta,$$

$$\alpha z_i^{\lambda} N_i^{\alpha\lambda-1} = \theta,$$

$$N_i = \left[\frac{\theta}{\alpha z_i^{\lambda}}\right]^{1/(\alpha\lambda-1)}.$$

Subsitute for N_i into the aggregate resource constraints:

$$\theta = \left[\int \big[\alpha z_i^{\lambda}\big]^{1/(1-\alpha\lambda)}\, di\right]^{1-\alpha\lambda},$$

$$C = \int \left[\frac{\theta}{\alpha z_i^{\lambda}}\right]^{\alpha/(\alpha\lambda-1)} di$$

$$= \left[\int z_i^{\lambda/(1-\alpha\lambda)}\, di\right]^{-\alpha} \int z_i^{\alpha\lambda/(1-\alpha\lambda)}\, di,$$

$$N_i = \left[\frac{\theta}{\alpha z_i^{\lambda}}\right]^{1/(\alpha\lambda-1)}$$

$$= \left[\frac{\alpha z_i^{\lambda}}{\left[\int \big[\alpha z_i^{\lambda}\big]^{1/(1-\alpha\lambda)}\, di\right]^{1-\alpha\lambda}}\right]^{1/(1-\alpha\lambda)}$$

$$= \frac{z_i^{\lambda/(1-\alpha\lambda)}}{\int z_i^{\lambda/(1-\alpha\lambda)}\, di},$$

$$\begin{aligned} A_i &= z_i^{\lambda} N_i^{\alpha\lambda-\alpha} \\ &= \frac{z_i^{(\lambda-\alpha\lambda)/(1-\alpha\lambda)}}{\left[\int z_i^{\lambda/(1-\alpha\lambda)}\, di\right]^{\alpha\lambda-\alpha}}. \end{aligned}$$

These expressions yield unique allocations. But do these allocations correspond to an equilibrium?

If $\alpha\lambda > 1$, then the formula for employment says it is inversely related to productivity. The reason is that the wage is identical across inhabited locations. With increasing returns the only way this can happen is if low productivity locations have higher TFP through the externality. If agents always believe this distribution of TFP, then the equilibrium is maintained. But there is another set of beliefs that yields higher utility. In particular, that everyone assumes TFP is zero everywhere except at the highest productivity shock cities. Everyone would locate there and there would be no incentive to move elsewhere—wages are zero everywhere else and coordinating on another set of beliefs lowers utility.

With $\alpha\lambda < 1$, this is not an issue. The household will never leave a location unoccupied because the marginal product of labor goes to infinity as labor goes to zero. Given the distribution of TFP implied by the FONCs, the program has a unique solution that implies the same distribution of TFP. There is only one distribution of TFP that solves the FONCs.

H.2. *Model Without Housing and With One-Period Time-to-Build Infrastructure*

We have

$$\begin{aligned} &\max \sum \beta^t \ln C_t \\ &\quad + \pi\left[\int A_{it} N_{it}^{\alpha} L_{it}^{1-\alpha}\, di - C - \int \left[K_{it+1} - (1-\kappa)K_{it}\right] di\right] \\ &\quad + \pi \int \varphi_i \left[K_{it}^{s} - L_{it}\right] di \\ &\quad + \pi\theta\left[1 - \int N_i\, di\right]. \end{aligned}$$

The necessary and sufficient conditions for an interior solution to the program for given A_i are

$$\begin{aligned} &A_{it}\alpha N_{it}^{\alpha-1} L_{it}^{1-\alpha} = \theta, \\ &A_{it}(1-\alpha) N_{it}^{\alpha} L_{it}^{-\alpha} = \varphi_{it}, \end{aligned}$$

$$K_{it}^{\varsigma} = L_{it},$$
$$1 = \beta E_{it}\big[\varphi_{it+1}\varsigma K_{it+1}^{\varsigma-1} + (1-\kappa)\big],$$
$$1 = \int N_{it}\,di,$$
$$C + \int \big[K_{it+1} - (1-\kappa)K_{it}\big]\,di = \int A_{it}N_{it}^{\alpha}L_{it}^{1-\alpha}\,di,$$
$$\frac{1}{C} = \pi.$$

Solve for A_i:

$$\begin{aligned} A_{it} &= z_{it}\big[A_{it}N_{it}^{\alpha}K_{it}^{\varsigma(1-\alpha)-\varsigma}\big]^{1-1/\lambda} \\ &= z_{it}^{\lambda}N_{it}^{\alpha\lambda-\alpha}K_{it}^{-\lambda\varsigma\alpha+\varsigma\alpha}. \end{aligned}$$

Substitute for A_i in labor FONC and solve for N_{it},

$$z_{it}^{\lambda}N_{it}^{\alpha\lambda-\alpha}K_{it}^{-\lambda\varsigma\alpha+\varsigma\alpha}\alpha N_{it}^{\alpha-1}K_{it}^{\varsigma(1-\alpha)} = \theta,$$
$$\alpha z_{it}^{\lambda}N_{it}^{\alpha\lambda-1}K_{it}^{\varsigma(1-\lambda\alpha)} = \theta,$$
$$\begin{aligned} N_{it} &= \left[\frac{\theta}{\alpha z_{it}^{\lambda}K_{it}^{\varsigma(1-\lambda\alpha)}}\right]^{1/(\alpha\lambda-1)} \\ &= \left[\frac{\theta}{\alpha z_{it}^{\lambda}}\right]^{1/(\alpha\lambda-1)} K_{it}^{\varsigma}, \end{aligned}$$

and in land FONC and solve for φ_{it},

$$z_{it}^{\lambda}N_{it}^{\alpha\lambda-\alpha}K_{it}^{-\lambda\varsigma\alpha+\varsigma\alpha}(1-\alpha)N_{it}^{\alpha}K_{it}^{-\varsigma\alpha} = \varphi_{it},$$
$$z_{it}^{\lambda}\left[\frac{\theta}{\alpha z_{it}^{\lambda}}\right]^{\alpha\lambda/(\alpha\lambda-1)}(1-\alpha) = \varphi_{it},$$
$$\varphi_{it} = (1-\alpha)z_{it}^{\lambda/(1-\alpha\lambda)}\big[\alpha\theta^{-1}\big]^{\alpha\lambda/(1-\alpha\lambda)},$$

which can be substituted into the infrastructure FONC as

$$1 = \beta E_{it}\big[(1-\alpha)z_{it+1}^{\lambda/(1-\alpha\lambda)}\big[\alpha\theta^{-1}\big]^{\alpha\lambda/(1-\alpha\lambda)}\varsigma K_{it+1}^{\varsigma-1} + (1-\kappa)\big],$$

so that

$$\beta(1-\alpha)\varsigma K_{it+1}^{\varsigma-1}E_{it}z_{it+1}^{\lambda/(1-\alpha\lambda)}\big[\alpha\theta^{-1}\big]^{\alpha\lambda/(1-\alpha\lambda)} = 1 - \beta(1-\kappa),$$

which yields

$$K_{it+1} = \left[\frac{1-\beta(1-\kappa)}{\beta(1-\alpha)\varsigma E_{it} z_{it+1}^{\lambda/(1-\alpha\lambda)}\left[\alpha\theta^{-1}\right]^{\alpha\lambda/(1-\alpha\lambda)}}\right]^{1/(\varsigma-1)}.$$

Substituting this expression for K_{it+1} (backed up one period) into the aggregate employment constraint yields

$$\begin{aligned}
1 &= \int \left[\frac{\theta}{\alpha z_{it}^{\lambda}}\right]^{1/(\alpha\lambda-1)} K_{it}^{\varsigma}\, di \\
&= \int \left[\frac{\theta}{\alpha z_{it}^{\lambda}}\right]^{1/(\alpha\lambda-1)} \left[\frac{1-\beta(1-\kappa)}{\beta(1-\alpha)\varsigma E_{it-1} z_{it}^{\lambda/(1-\alpha\lambda)}\left[\alpha\theta^{-1}\right]^{\alpha\lambda/(1-\alpha\lambda)}}\right]^{\varsigma/(\varsigma-1)} di, \\
\theta &= \left[\int \left[\frac{1}{\alpha z_{it}^{\lambda}}\right]^{1/(\alpha\lambda-1)} \right. \\
&\quad \left. \times \left[\frac{1-\beta(1-\kappa)}{\beta(1-\alpha)\varsigma\alpha^{\alpha\lambda/(1-\alpha\lambda)} E_{it-1} z_{it}^{\lambda/(1-\alpha\lambda)}}\right]^{\varsigma/(\varsigma-1)} di\right]^{(\alpha\lambda-1)(\varsigma-1)/(1-\varsigma+\alpha\lambda\varsigma)}.
\end{aligned}$$

From this we obtain

$$\begin{aligned}
\theta &= f\left(z_{it-1}, E_{it-1} z_{it}^{a_f} : i \in [0,1]\right) \\
&= f\left(z_{it}, E_{it} z_{it+1}^{a_f} : i \in [0,1]\right),
\end{aligned}$$

where a_f is a scalar defined by algebraic expressions of the underlying model parameters. Similarly,

$$\begin{aligned}
K_{it+1} &= g\left(z_{it}, E_{it} z_{it+1}^{a_g} : i \in [0,1]\right), \\
K_{it} &= g\left(z_{it-1}, E_{it-1} z_{it}^{a_g} : i \in [0,1]\right).
\end{aligned}$$

Here the two expressions for θ reflect that we are calculating the stationary solution, so it does not matter which date we calculate it. The second expression for the local capital stock merely backs up the first one—in a stationary equilibrium both will be satisfied.

Note that in steady state,

$$\int K_{it+1}\, di = \int K_{it}\, di.$$

Therefore,

$$\begin{aligned}
\int \left[K_{it+1} - (1-\kappa)K_{it}\right] di &= \int K_{it+1} - \int K_{it} + \kappa \int K_{it} \\
&= \kappa \int K_{it}.
\end{aligned}$$

With this fact, the aggregate final good constraint becomes

$$\begin{aligned} C+\kappa\int K_{it}\,di &= \int A_{it}N_{it}^{\alpha}K_{it}^{\varsigma(1-\alpha)}\,di \\ &= \int z_{it}^{\lambda}N_{it}^{\alpha\lambda}K_{it}^{\varsigma(1-\alpha)\lambda}\,di \\ &= \int z_{it}^{\lambda}\left[\frac{\theta}{\alpha z_{it}^{\lambda}K_{it}^{\varsigma(1-\alpha)\lambda}}\right]^{\alpha\lambda/(\alpha\lambda-1)}K_{it}^{\varsigma(1-\alpha)\lambda}\,di \\ &= \int z_{it}^{\lambda}\left[\frac{\theta}{\alpha z_{it}^{\lambda}}\right]^{\alpha\lambda/(\alpha\lambda-1)}K_{it}^{\varsigma(1-\alpha)\lambda\alpha\lambda/(1-\alpha\lambda)}K_{it}^{\varsigma(1-\alpha)\lambda}\,di, \end{aligned}$$

which using the expressions for θ and K_{it} yields

$$C = h\big(z_{it-1}, E_{it-1}z_{it}^{a_h} : i\in[0,1]\big).$$

Finally, after substituting for θ and K_{it} in an earlier derivation, we have

$$N_{it} = m\big(z_{it}, \{z_{it-1}, E_{it-1}z_{it}^{a_m} : i\in[0,1]\}\big).$$

With a finite dimensional Markov chain (satisfying usual regularity conditions) for z_i, a nontrivial solution exists and is unique if $\alpha\lambda<1$. This is seen by considering the marginal product of labor

$$\alpha z_{it}^{\lambda}N_{it}^{\alpha\lambda-1}K_{it}^{\varsigma(1-\lambda\alpha)},$$

which indicates that unless the condition is satisfied, the marginal product of labor goes to zero as employment does. In this case, we cannot rule out corner solutions for at least some locations. If they are all occupied, then the above expressions can be used to solve for the unique equilibrium.

We now extend the model to imperfect substitutes.

H.3. *Model Without Housing and With Infrastructure and Imperfect Substitutes*

We have

$$\begin{aligned} &\max \sum \beta^t \ln C_t \\ &\quad + \pi_t\left[\left(\int Y_{it}^{\eta}\,di\right)^{1/\eta} di - C - \int\big[K_{it+1}-(1-\kappa)K_{it}\big]\,di\right] \\ &\quad + \pi\int \varphi_{it}\big[K_{it}^{\varsigma}-L_{it}\big]\,di + \pi\int q_{it}\big[A_{it}N_{it}^{\alpha}L_{it}^{1-\alpha}-Y_{it}\big]\,di \\ &\quad + \pi_t\theta_t\left[1-\int N_{it}\,di\right]. \end{aligned}$$

The necessary and sufficient conditions for an interior solution to the program for given A_i are

$$q_{it}A_{it}\alpha N_{it}^{\alpha-1}L_{it}^{1-\alpha}=\theta,$$
$$q_{it}A_{it}(1-\alpha)N_{it}^{\alpha}L_{it}^{-\alpha}=\varphi_{it},$$
$$q_{it}=Y^{1-\eta}Y_{it}^{\eta-1},$$
$$K_{it}^{\varsigma}=L_{it},$$
$$1=\beta E_{it}\left[\varphi_{it+1}\varsigma K_{it+1}^{\varsigma-1}+(1-\kappa)\right],$$
$$1=\int N_{it}\,di,$$
$$C+\int\left[K_{it+1}-(1-\kappa)K_{it}\right]di=\left[\int\left[A_{it}N_{it}^{\alpha}L_{it}^{1-\alpha}\right]^{\eta}di\right]^{1/\eta},$$
$$\frac{1}{C}=\pi.$$

Solve for

$$\begin{aligned}A_{it}&=z_{it}\left[A_{it}N_{it}^{\alpha}K_{it}^{\varsigma(1-\alpha)-\varsigma}\right]^{1-1/\lambda}\\&=z_{it}^{\lambda}N_{it}^{\alpha\lambda-\alpha}K_{it}^{-\lambda\varsigma\alpha+\varsigma\alpha}\end{aligned}$$

and

$$\begin{aligned}q_{it}&=Y^{1-\eta}Y_{it}^{\eta-1}\\&=Y^{1-\eta}\left[A_{it}N_{it}^{\alpha}L_{it}^{1-\alpha}\right]^{\eta-1}\\&=Y^{1-\eta}\left[z_{it}^{\lambda}N_{it}^{\alpha\lambda}K_{it}^{\varsigma(1-\lambda\alpha)}\right]^{\eta-1}.\end{aligned}$$

Substitute in labor FONC

$$Y^{1-\eta}\left[z_{it}^{\lambda}N_{it}^{\alpha\lambda}K_{it}^{\varsigma(1-\lambda\alpha)}\right]^{\eta-1}z_{it}^{\lambda}N_{it}^{\alpha\lambda-\alpha}K_{it}^{-\lambda\varsigma\alpha+\varsigma\alpha}\alpha N_{it}^{\alpha-1}K_{it}^{\varsigma-\varsigma\alpha}=\theta,$$

which yields

$$N_{it}=\left[\frac{\theta}{Y^{1-\eta}z_{it}^{\lambda\eta}\alpha}\right]^{1/(\eta\alpha\lambda-1)}K_{it}^{\eta\varsigma(1-\lambda\alpha)/(1-\eta\alpha\lambda)}.$$

Substitute in land FONC

$$Y^{1-\eta}\left[z_{it}^{\lambda}N_{it}^{\alpha\lambda}K_{it}^{\varsigma(1-\lambda\alpha)}\right]^{\eta-1}z_{it}^{\lambda}N_{it}^{\alpha\lambda}K_{it}^{-\lambda\varsigma\alpha}(1-\alpha)=\varphi_{it},$$
$$Y^{1-\eta}K_{it}^{\varsigma(1-\lambda\alpha)(\eta-1)}z_{it}^{\eta\lambda}N_{it}^{\eta\alpha\lambda}K_{it}^{-\lambda\varsigma\alpha}(1-\alpha)=\varphi_{it}.$$

Simplify exponent on K_{it},

$$K_{it}^{\varsigma(1-\lambda\alpha)(\eta-1)}K_{it}^{-\lambda\varsigma\alpha} = K_{it}^{\varsigma(1-\lambda\alpha)\eta-\varsigma(1-\lambda\alpha)-\lambda\varsigma\alpha} = K_{it}^{\varsigma(1-\lambda\alpha)\eta-\varsigma},$$

so that

$$\varphi_{it} = Y^{1-\eta}z_{it}^{\eta\lambda}N_{it}^{\eta\alpha\lambda}K_{it}^{\varsigma(1-\lambda\alpha)\eta-\varsigma}(1-\alpha) = Y^{1-\eta}z_{it}^{\eta\lambda}\left[\frac{\theta}{Y^{1-\eta}z_{it}^{\lambda\eta}\alpha}\right]^{\eta\alpha\lambda/(\eta\alpha\lambda-1)} \times K_{it}^{\eta\alpha\lambda\eta\varsigma(1-\lambda\alpha)/(1-\eta\alpha\lambda)}K_{it}^{\varsigma(1-\lambda\alpha)\eta-\varsigma}(1-\alpha).$$

Simplify the exponent on K_{it},

$$\frac{\eta\alpha\lambda\eta\varsigma(1-\lambda\alpha)}{1-\eta\alpha\lambda}+\varsigma(1-\lambda\alpha)\eta-\varsigma = \frac{\eta\alpha\lambda\eta\varsigma(1-\lambda\alpha)+\varsigma(1-\lambda\alpha)\eta-\varsigma-\eta\alpha\lambda(\varsigma(1-\lambda\alpha)\eta-\varsigma)}{1-\eta\alpha\lambda} = \frac{\eta\alpha\lambda\eta\varsigma-\lambda\alpha\eta\alpha\lambda\eta\varsigma+\varsigma(1-\lambda\alpha)\eta-\varsigma-\eta\alpha\lambda(\varsigma\eta-\lambda\alpha\varsigma\eta-\varsigma)}{1-\eta\alpha\lambda} = \frac{\varsigma\eta-\varsigma}{1-\eta\alpha\lambda},$$

so that

$$\varphi_{it} = Y^{1-\eta}z_{it}^{\eta\lambda/(1-\eta\alpha\lambda)}\left[\frac{\theta}{Y^{1-\eta}\alpha}\right]^{\eta\alpha\lambda/(\eta\alpha\lambda-1)}K_{it}^{(\varsigma\eta-\varsigma)/(1-\eta\alpha\lambda)}(1-\alpha),$$

which can be substituted into the infrastructure FONC as

$$1 = \beta E_{it}\big[\varphi_{it+1}\varsigma K_{it+1}^{\varsigma-1}+(1-\kappa)\big] = \beta E_{it}\left[Y^{1-\eta}z_{it+1}^{\eta\lambda/(1-\eta\alpha\lambda)}\left[\frac{\theta}{Y^{1-\eta}\alpha}\right]^{\eta\alpha\lambda/(\eta\alpha\lambda-1)} \times K_{it+1}^{(\varsigma\eta-1-\eta\alpha\lambda\varsigma+\eta\alpha\lambda)/(1-\eta\alpha\lambda)}\varsigma(1-\alpha)+(1-\kappa)\right]$$

to yield

$$K_{it+1} = \Big[\big(1-\beta(1-\kappa)\big) \Big/\Big(\theta^{\eta\alpha\lambda/(\eta\alpha\lambda-1)}(1-\alpha)\varsigma\beta Y^{(1-\eta)/(1-\eta\alpha\lambda)} \times E_{it}z_{it+1}^{\eta\lambda/(1-\eta\alpha\lambda)}\Big)\Big]^{(1-\eta\alpha\lambda)/(\varsigma\eta-1-\eta\alpha\lambda\varsigma+\eta\alpha\lambda)}.$$

Solve for Y_{it} and substitute into the expression for Y,

$$\begin{aligned}
Y_{it} &= A_{it} N_{it}^{\alpha} L_{it}^{1-\alpha} \\
&= z_{it}^{\lambda} N_{it}^{\alpha\lambda} K_{it}^{\varsigma(1-\lambda\alpha)} \\
&= z_{it}^{\lambda} \left[\frac{\theta}{Y^{1-\eta} z_{it}^{\lambda\eta} \alpha} \right]^{\alpha\lambda/(\eta\alpha\lambda-1)} K_{it}^{\alpha\lambda\eta\varsigma(1-\lambda\alpha)/(1-\eta\alpha\lambda)} K_{it}^{\varsigma(1-\lambda\alpha)} \\
&= z_{it}^{\lambda} \left[\frac{\theta}{Y^{1-\eta} z_{it}^{\lambda\eta} \alpha} \right]^{\alpha\lambda/(\eta\alpha\lambda-1)} K_{it}^{\varsigma(1-\lambda\alpha)/(1-\eta\alpha\lambda)} \\
&= z_{it}^{\lambda} \left[\frac{\theta}{Y^{1-\eta} z_{it}^{\lambda\eta} \alpha} \right]^{\alpha\lambda/(\eta\alpha\lambda-1)} \\
&\quad \times \Big[\big(1 - \beta(1-\kappa)\big) \\
&\quad / \big(\beta Y^{(1-\eta)/(1-\eta\alpha\lambda)} \theta^{\eta\alpha\lambda/(\eta\alpha\lambda-1)} (1-\alpha)\varsigma \\
&\quad \times E_{it-1} z_{it+1}^{\eta\lambda/(1-\eta\alpha\lambda)} \big) \Big]^{\varsigma(1-\lambda\alpha)/(\varsigma\eta-1-\eta\alpha\lambda\varsigma+\eta\alpha\lambda)}.
\end{aligned}$$

Substituting into the definition of Y,

$$\begin{aligned}
Y = \Bigg[\int & z_{it}^{\eta\lambda} \left[\frac{\theta}{Y^{1-\eta} z_{it}^{\lambda\eta} \alpha} \right]^{\eta\alpha\lambda/(\eta\alpha\lambda-1)} \\
& \times \Big[\big(1 - \beta(1-\kappa)\big) \\
& / \big(\beta Y^{(1-\eta)/(1-\eta\alpha\lambda)} \theta^{\eta\alpha\lambda/(\eta\alpha\lambda-1)} (1-\alpha)\varsigma \\
& \times E_{it-1} z_{it+1}^{\eta\lambda/(1-\eta\alpha\lambda)} \big) \Big]^{\varsigma(1-\lambda\alpha)/(\varsigma\eta-1-\eta\alpha\lambda\varsigma+\eta\alpha\lambda)} \, di \Bigg]^{1/\eta}.
\end{aligned}$$

Using capital letters to denote complicated expressions of underlying parameters, we arrive at

$$Y = \Gamma \theta^{B} Y^{D} \left[\int z_{it}^{A} \left[E_{it-1} z_{it}^{\eta\lambda/(1-\eta\alpha\lambda)} \right]^{\varsigma(1-\lambda\alpha)/(\varsigma\eta-1-\eta\alpha\lambda\varsigma+\eta\alpha\lambda)} \, di \right]^{1/\eta},$$

$$Y = \left(\Gamma \theta^{B} \left[\int z_{it}^{A} \left[E_{it-1} z_{it}^{\eta\lambda/(1-\eta\alpha\lambda)} \right]^{\varsigma(1-\lambda\alpha)/(\varsigma\eta-1-\eta\alpha\lambda\varsigma+\eta\alpha\lambda)} \, di \right]^{1/\eta} \right)^{1/(1-D)}.$$

Substitute for variables in the aggregate employment constraint:

$$\begin{aligned}
1 &= \int N_i \, di \\
&= \int \left[\frac{\theta}{Y^{1-\eta} z_{it}^{\lambda\eta} \alpha} \right]^{1/(\eta\alpha\lambda-1)} K_{it}^{\eta\varsigma(1-\lambda\alpha)/(1-\eta\alpha\lambda)} \, di
\end{aligned}$$

$$= \int \left[\frac{\theta}{Y^{1-\eta} z_{it}^{\lambda\eta} \alpha} \right]^{1/(\eta\alpha\lambda-1)}$$
$$\times \left[\left(1 - \beta(1-\kappa)\right) \right.$$
$$/ \left(\theta^{\eta\alpha\lambda/(\eta\alpha\lambda-1)} (1-\alpha) \varsigma\beta Y^{(1-\eta)/(1-\eta\alpha\lambda)} \right.$$
$$\left. \left. \times E_{it-1} z_{it}^{\eta\lambda/(1-\eta\alpha\lambda)} \right) \right]^{\eta\varsigma(1-\lambda\alpha)/(\varsigma\eta-1-\eta\alpha\lambda\varsigma+\eta\alpha\lambda)} di.$$

After substituting for Y, we arrive at

$$\theta = f\left(z_{it-1}, E_{it-1} z_{it}^{a_f} : i \in [0,1]\right),$$
$$Y = l\left(z_{it-1}, E_{it-1} z_{it}^{a_l} : i \in [0,1]\right),$$
$$K_{it+1} = g\left(z_{it}, E_{it} z_{it}^{a_g} : i \in [0,1]\right),$$
$$N_{it} = m\left(z_{it}, \left\{z_{it-1}, E_{it-1} z_{it}^{a_m} : i \in [0,1]\right\}\right),$$
$$C + \kappa \int g\left(z_{it-1}, E_{it-1} z_{it}^{a_g} : i \in [0,1]\right) di$$
$$= l\left(z_{it-1}, E_{it-1} z_{it}^{a_l} : i \in [0,1]\right),$$
$$C = h\left(z_{it-1}, E_{it-1} z_{it}^{a_g}, E_{it-1} z_{it}^{a_l} : i \in [0,1]\right).$$

Notice that we can use the expression for

$$N_{it} = \left[\frac{\theta}{Y^{1-\eta} z_{it}^{\lambda\eta} \alpha} \right]^{1/(\eta\alpha\lambda-1)} K_{it}^{\eta\varsigma(1-\lambda\alpha)/(1-\eta\alpha\lambda)}$$

to verify that $\eta\alpha\lambda < 1$ is not sufficient to guarantee employment is not inversely related to productivity. For $\eta < 0$, we would have an inverse relationship. This arises because high enough complementarity deters allocating workers to locations because output at those locations would be too high relative to other, complementary, locations.

The marginal product of labor is given by

$$\alpha Y^{1-\eta} z_{it}^{\lambda\eta} N_{it}^{\eta\alpha\lambda-1} K_{it}^{\eta\varsigma(1-\lambda\alpha)}.$$

Clearly if $\eta\alpha\lambda > 1$, then as employment goes to zero, then so does the marginal product of labor. Workers would be allocated to the city with the highest productivity (for any given infrastructure stock). We require $\eta\alpha\lambda < 1$ to guarantee that all locations are occupied. If they are all occupied, then the above expressions can be used to solve for the unique equilibrium.

H.4. *Model Without Housing and With Infrastructure, Imperfect Substitutes, and Freely Mobile Equipment*

We have

$$\begin{aligned}
&\max \sum \beta^t \ln C_t \\
&\quad + \pi_t \bigg[\left(\int Y_{it}^{\eta}\, di \right)^{1/\eta} - C - \int \big[D_{it+1} - (1-\kappa_d) D_{it} \big]\, di \\
&\quad - \big[K_{t+1} - (1-\kappa_k) K_t \big] \bigg] \\
&\quad + \pi \varphi_{it} \big[D_{it}^{\varsigma} - L_{it} \big] \\
&\quad + \pi q_{it} \big[A_{it} K_{it}^{\phi\alpha} N_{it}^{\phi(1-\alpha)} L_{it}^{(1-\phi)} - Y_{it} \big] \\
&\quad + \pi_t \theta_t \bigg[1 - \int N_{it}\, di \bigg] \\
&\quad + \pi_t r_t \bigg[1 - \int D_{it}\, di \bigg].
\end{aligned}$$

The necessary and sufficient conditions for an interior solution to the program for given A_i are

$$\begin{aligned}
&q_{it} A_{it} \phi (1-\alpha) K_{it}^{\phi\alpha} N_{it}^{\phi(1-\alpha)-1} L_{it}^{1-\phi} = \theta, \\
&q_{it} A_{it} (1-\phi) K_{it}^{\phi\alpha} N_{it}^{(1-\alpha)\phi} L_{it}^{-\phi} = \varphi_{it}, \\
&q_{it} A_{it} \phi\alpha K_{it}^{\phi\alpha-1} N_{it}^{\phi(1-\alpha)} L_{it}^{1-\phi} = r, \\
&q_{it} = Y^{1-\eta} Y_{it}^{\eta-1}, \\
&D_{it}^{\varsigma} = L_{it}, \\
&1 = \beta E_{it} \big[\varphi_{it+1} \varsigma D_{it+1}^{\varsigma-1} + (1-\kappa_d) \big], \\
&1 = \beta [r + 1 - \kappa_k], \\
&1 = \int N_{it}\, di, \\
&K = \int K_{it}\, di, \\
&C + \int \big[D_{it+1} - (1-\kappa_d) D_{it} \big]\, di + K_{t+1} - (1-\kappa_k) K_{it} = \big[Y_{it}^{\eta}\, di \big]^{1/\eta}, \\
&A_{it} K_{it}^{\phi\alpha} N_{it}^{\phi(1-\alpha)} L_{it}^{1-\phi} = Y_{it},
\end{aligned}$$

$$\frac{1}{C} = \pi.$$

Solve for K_{it}, noting that the rental rate on equipment r is determined as a function of parameters by the intertemporal condition:

$$\frac{q_{it}A_{it}\phi(1-\alpha)K_{it}^{\phi\alpha}N_{it}^{\phi(1-\alpha)-1}L_{it}^{1-\phi}}{q_{it}A_{it}\phi\alpha K_{it}^{\phi\alpha-1}N_{it}^{\phi(1-\alpha)}L_{it}^{1-\phi}} = \frac{\theta}{r},$$

$$\frac{(1-\alpha)K_{it}}{\alpha N_{it}} = \frac{\theta}{r},$$

$$K_{it} = \frac{\alpha\theta}{(1-\alpha)r}N_{it}.$$

After substituting into land and labor FONCs and expressions for Y_{it}, A_{it}, and q_{it}, the problem is essentially the same as above except for the additional term in the aggregate equipment and resource constraints. Notice from the aggregate equipment constraint that

$$\begin{aligned} K &= \int K_{it}\,di \\ &= \int \frac{\alpha\theta}{(1-\alpha)r}N_{it}\,di \\ &= \frac{\alpha\theta}{(1-\alpha)r}. \end{aligned}$$

After making the substitution for K in the aggregate resource constraint and following the steps in the model without equipment, we have a solution.

It is helpful to do the early parts of the derivation so we can find the restriction on parameters necessary for a well behaved interior solution.

Solve for

$$\begin{aligned} A_{it} &= z_{it}^{\phi(1-\alpha)}\left[A_{it}K_{it}^{\phi\alpha}N_{it}^{\phi(1-\alpha)}L_{it}^{1-\phi}/L_{it}\right]^{1-1/\lambda} \\ &= z_{it}^{\phi(1-\alpha)}\left[A_{it}\left(\frac{\alpha\theta}{(1-\alpha)r}N_{it}\right)^{\phi\alpha}N_{it}^{\phi(1-\alpha)}D_{it}^{\varsigma(1-\phi)-\varsigma}\right]^{1-1/\lambda} \\ &= z_{it}^{\lambda\phi(1-\alpha)}\left(\frac{\alpha\theta}{(1-\alpha)r}\right)^{\phi\alpha(\lambda-1)}N_{it}^{\phi\lambda-\phi}D_{it}^{-\lambda\varsigma\phi+\varsigma\phi} \end{aligned}$$

and

$$\begin{aligned} q_{it} &= Y^{1-\eta}Y_{it}^{\eta-1} \\ &= Y^{1-\eta}\left[A_{it}K_{it}^{\phi\alpha}N_{it}^{\phi(1-\alpha)}L_{it}^{1-\phi}\right]^{\eta-1} \end{aligned}$$

$$= Y^{1-\eta}\left[z_{it}^{\lambda\phi(1-\alpha)}\left(\frac{\alpha\theta}{(1-\alpha)r}\right)^{\phi\alpha(\lambda-1)}N_{it}^{\phi\lambda-\phi}D_{it}^{-\lambda\varsigma\phi+\varsigma\phi}\right.$$
$$\left.\times\left(\frac{\alpha\theta}{(1-\alpha)r}\right)^{\phi\alpha}N_{it}^{\phi}D_{it}^{\varsigma(1-\phi)}\right]^{\eta-1}$$
$$= Y^{1-\eta}\left[z_{it}^{\lambda\phi(1-\alpha)}\left(\frac{\alpha\theta}{(1-\alpha)r}\right)^{\phi\alpha\lambda}N_{it}^{\phi\lambda}D_{it}^{\varsigma-\lambda\varsigma\phi}\right]^{\eta-1}.$$

Substitute in the labor FONC,

$$q_{it}A_{it}\phi(1-\alpha)K_{it}^{\phi\alpha}N_{it}^{\phi(1-\alpha)-1}L_{it}^{1-\phi}=\theta,$$
$$\phi(1-\alpha)Y^{1-\eta}\left[z_{it}^{\lambda\phi(1-\alpha)}\left(\frac{\alpha\theta}{(1-\alpha)r}\right)^{\phi\alpha\lambda}N_{it}^{\phi\lambda}D_{it}^{\varsigma-\lambda\varsigma\phi}\right]^{\eta-1}z_{it}^{\lambda\phi(1-\alpha)}$$
$$\times\left(\frac{\alpha\theta}{(1-\alpha)r}\right)^{\phi\alpha(\lambda-1)}N_{it}^{\phi\lambda-\phi}D_{it}^{-\lambda\varsigma\phi+\varsigma\phi}\left(\frac{\alpha\theta}{(1-\alpha)r}\right)^{\phi\alpha}N_{it}^{\phi-1}L_{it}^{1-\phi}=\theta,$$
$$z_{it}^{\eta\lambda\phi(1-\alpha)}\left(\frac{\alpha\theta}{(1-\alpha)r}\right)^{\eta\phi\alpha\lambda}\phi(1-\alpha)Y^{1-\eta}N_{it}^{\eta\phi\lambda-1}D_{it}^{\eta\varsigma-\eta\lambda\varsigma\phi}=\theta,$$

so that we still have the necessary condition $\eta\phi\lambda < 1$ to avoid corner solutions.

APPENDIX I: ESTIMATION ROBUSTNESS

We first confirm that making decisions earlier requires additional lags for the instruments. Suppose firms' choices are made at $t-T$, $T>0$. We focus on the model without housing and capital, and make use of log linear approximations. Note that profits equal zero only in expectation, not by date and state. Here, to make the notation simpler, we index cities by their location on the unit interval rather than productivity history as we do in the paper and elsewhere in this document.

We use the following implication of a conditional expectation. For any variable x,

$$x_{it}=E_{it-T}x_{it}+\sum_{j=0}^{T-1}u_{it-j},$$

where

$$u_{it-j}\perp\Omega_{t-j-1}$$

denotes one-step-ahead forecast errors.

In the competitive equilibrium firms, set labor as

$$E_{it-T}\frac{\alpha Y_{it}}{N_{it,t-T}} = E_{it-T}W_{it},$$

where $N_{it,t-1}$ denotes employment for t chosen at $t-T$. Log linearizing yields

$$E_{it-T}e^{\ln\alpha+\ln Y_{it}-\ln N_{it,t-T}} = E_{it-T}e^{\ln W_{it}},$$
$$n_{it,t-T} = E_{it-T}y_{it} - E_{it-T}w_{it},$$

where

$$x_{it} = \ln X_{it} - E\ln X_{it}$$

for variable X. Similarly, log linearizing the land FONC yields

$$l_{bit,t-T} = E_{it-T}y_{it} - E_{it-T}r_{it}.$$

The TFP is

$$a_{it} = y_{it} - (1-\alpha)l_{bit,t-T} + \alpha n_{it,t-T}.$$

Using the implication of the conditional expectation gives

$$n_{it,t-T} = y_{it} - w_{it} + \sum_{j=0}^{T-1} u_{it-j},$$

$$l_{bit,t-T} = y_{it} - r_{it} + \sum_{j=0}^{T-1} v_{it-j}.$$

Therefore,

$$a_{it} = (1-\alpha)r_{it} + \alpha w_{it} + \sum_{j=0}^{T-1} u_{it-j} + \sum_{j=0}^{T-1} v_{it-j}.$$

The density externality is written as

$$y_{it} - l_{bi,t-T} = r_{it} + \sum_{j=0}^{T-1} v_{it-j}.$$

Clearly TFP and the externality are correlated with variables lagged up to $T-1$ times. Therefore, valid instruments must be lagged at least T times.

A similar result holds for the case when non-infrastructure capital is allocated one period in advance of the TFP shock. In this case, non-infrastructure investment implies a first order condition similar to that for infrastructure in the baseline version of the model. Using the notation from the first section of this document,

$$\pi_t P_{bt} = \beta \pi_{t+1} \sum_{z_{t+1}} \left[r_{bt+1}\left(z^{t+1}\right) + P_{bt+1}(1-\kappa_b) \right] Q(z_t, z_{t+1}).$$

Using the fact that $\pi_t = \pi_{t+1}$ yields

$$\sum_{z_{t+1}} r_{bt+1}\left(z^{t+1}\right) Q(z_t, z_{t+1}) = P_{bt}/\beta - P_{bt+1}(1-\kappa_b) = r_{bt+1},$$

where r_{bt+1} denotes the unconditional average rent on non-infrastructure capital. By the definition of a conditional expectation,

$$r_{bt+1}\left(z^{t+1}\right) = \sum_{z_{t+1}} r_{bt+1}\left(z^{t+1}\right) Q(z_t, z_{t+1}) + e(\varepsilon_{t+1}) = r_{bt+1} + e(\varepsilon_{t+1}),$$

where $\varepsilon_{t+1} \perp z_t$ is the TFP shock. It follows that the right hand side of the measurement equation for city-specific TFP, equation (29) in the main text, is identical except for the addition of an error term orthogonal to earlier dated variables. Our instrumental variables strategy remains valid.

Now we consider several perturbations to the estimation that involve when the density externality impacts TFP and the timing of allocation decision. We allow TFP to depend on lagged density, and averages of current and lagged density. This involves replacing r_{it} in the main text's equation (26) with lagged values or averages of current and lagged values. The following table reports our findings:

Perturbation of Estimation	δ	se	Λ	se
Replace r_{it} with r_{it-1}	1.040	0.024	0.101	0.052
Replace r_{it} with r_{it-2}	1.040	0.018	0.101	0.041
Replace r_{it} with $(r_{it} + r_{it-1})/2$	1.045	0.027	0.109	0.059
Replace r_{it} with $(r_{it-1} + r_{it-2})/2$	1.042	0.021	0.103	0.043
Replace r_{it} with $(r_{it} + r_{it-1} + r_{it-2})/3$	1.045	0.024	0.108	0.055

Suppose allocations are picked T periods in advance. As discussed above, this introduces expectational errors in addition to technology shocks in the main text's equation (26). To address this, we need to back up the instruments. Backing up the instruments of course lowers their predictive power. The following table reports our findings when we increase the lag on the instruments:

Perturbation of Estimation	δ	se	Λ	se
Replace instruments lagged 2 with 3	1.119	0.090	0.229	0.185
Replace instruments lagged 2 with 4	1.244	0.225	0.510	0.814

TABLE S.XVI

ESTIMATION ROBUSTNESS RESULTS

Perturbation of Estimation With Lagged 3 Instruments	δ	se	Λ	se
Replace r_{it} with r_{it-1}	1.072	0.037	0.152	0.085
Replace r_{it} with r_{it-2}	1.043	0.021	0.106	0.049
Replace r_{it} with $(r_{it}+r_{it-1})/2$	1.091	0.054	0.182	0.119
Replace r_{it} with $(r_{it-1}+r_{it-2})/2$	1.055	0.027	0.125	0.063
Replace r_{it} with $(r_{it}+r_{it-1}+r_{it-2})/3$	1.069	0.037	0.146	0.083
Perturbation of Estimation With Lagged 4 Instruments	δ	se	Λ	se
Replace r_{it} with r_{it-1}	1.091	0.059	0.202	0.161
Replace r_{it} with r_{it-2}	1.056	0.032	0.138	0.070
Replace r_{it} with $(r_{it}+r_{it-1})/2$	1.136	0.096	0.285	0.295
Replace r_{it} with $(r_{it-1}+r_{it-2})/2$	1.070	0.042	0.163	0.101
Replace r_{it} with $(r_{it}+r_{it-1}+r_{it-2})/3$	1.093	0.058	0.205	0.155

Table S.XVI reports our findings when we combine the two perturbations.

APPENDIX J: VERIFYING THE LUTTMER (2007) PROPERTY

In our context, the Luttmer property is that the ratio of the Zipf coefficient for technology relative to that for population is invariant to technology's Zipf coefficient, holding all other model parameters fixed. Therefore, for otherwise fixed parameters, adjusting γ_z to yield a different Zipf coefficient for technology should not change the Zipf coefficient ratio implied by the model if the model satisfies the Luttmer property.

We verify the Luttmer property in our model by holding all parameters at their estimated values except for η, λ, and γ_z. Recall that η, holding other parameters fixed, can be adjusted to change the Zipf ratio. We consider $\eta = 0.927$ and $\eta = 0.9$, and $\lambda = 1.02$ and $\lambda = 1.05$. For the four combinations of parameters, we vary γ_z so that the approximate technology Zipf coefficient equals -10, -19, -20, and -21, resolving the model and calculating the Zipf coefficient ratio for each case.

We face a challenge when approximating the technology process. In particular, holding fixed the domain of the approximation, adjusting the Zipf coefficient changes the mass of the technology on the largest grid points and this introduces a bias into our Zipf coefficient ratio calculation. Therefore, in our calculations, we adjust the domain of the grid so that the mass on the largest grid point is approximately the same.[13]

Table S.XVII reports our findings. If the Luttmer property holds, then the technology–population ratio of Zipf coefficients should be invariant to the

[13]The mass on the last grid point is 0.0060, 0.0060, 0.0060, and 0.0061 for Zipf coefficients -21, -20, -19, and -10.

TABLE S.XVII

VALIDATION OF THE LUTTMER PROPERTY

Technology's Zipf Coefficient	Ratio of Zipf Coefficients			
	$\lambda = 1.05$		$\lambda = 1.02$	
	$\eta = 0.927$	$\eta = 0.9$	$\eta = 0.927$	$\eta = 0.9$
−21	4.71	3.80	4.27	3.49
−20	4.71	3.80	4.27	3.49
−19	4.72	3.83	4.28	3.49
−10	4.69	3.83	4.26	3.49

technology Zipf coefficient for each particular η and λ combination. While this property holds "exactly" in only one case, the ratios are roughly constant for the other cases and the variation does not follow a pattern that would suggest significant departures from the Luttmer property. In this sense, we say that we verify numerically that the Luttmer property holds approximately in our model.

Recent results in Davis, Fisher, and Veracierto (2013) suggested that this finding is not unique to our specification. These authors study a similar model driven by a similar technology process, but without an agglomeration externality. They found that the Luttmer property holds and confirmed that the kind of model we study can be made consistent with empirical estimates of the level and the ratio of the Zipf coefficients for technology and population.

REFERENCES

DAVIS, M. A., AND J. H. HEATHCOTE (2007): "The Price and Quantity of Residential Land in the United States," *Journal of Monetary Economics*, 54 (8), 2595–2620. [25]

DAVIS, M., J. FISHER, AND M. VERACIERTO (2013): "Migration, Housing and Labor Reallocation," Unpublished Manuscript, Chicago Fed and University of Wisconsin. [65]

ERICKSON, T., AND T. WHITED (2002): "Two-Step GMM Estimation of the Errors-in-Variables Model Using High-Order Moments," *Econometric Theory*, 18 (3), 776–799. [46]

LUTTMER, E. (2007): "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, 122 (3), 1103–1144. [1,64]

NEWEY, W., AND D. MCFADDEN (1994): "Large Sample Estimation and Hypothesis Testing," in *The Handbook of Econometrics*, Vol. 4, ed. by R. Engle and D. McFadden. Amsterdam: North-Holland, 2111–2245. [45]

Dept. of Real Estate and Urban Land Economics, University of Wisconsin–Madison, Madison, WI 53706, U.S.A.; mdavis@bus.wisc.edu,

Federal Reserve Bank of Chicago, Chicago, IL 60637, U.S.A.; jfisher@frbchi.org,

and

Simon Business School, University of Rochester, Rochester, NY 14627, U.S.A.; toni.whited@simon.rochester.edu.

Manuscript received January, 2010; final revision received September, 2013.