Macroeconomic Implications of Agglomeration*

Morris A. Davis
University of Wisconsin
mdavis@bus.wisc.edu

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

Toni M. Whited
University of Rochester
toni.whited@simon.rochester.edu

December 30, 2009

Abstract

We construct a dynamic general equilibrium model of cities and use it to estimate the effect of local agglomeration on per capita consumption growth. Agglomeration affects growth through the density of economic activity: higher production per unit of land raises local productivity. Firms take productivity as given; produce using a technology that has constant returns in developed land, capital, and labor; and accumulate land and capital. If land prices are rising, as they are empirically, firms economize on land. This behavior increases density and contributes to growth. We use a panel of U.S. cities and our model's predicted relationship among wages, output prices, housing rents, and labor quality to estimate the net effect of agglomeration on local wages. The impact of agglomeration on the level of wages is estimated to be 2 percent. Combined with our model and observed increases in land prices, this estimate implies that agglomeration raises per capita consumption growth by 10 percent.

JEL Classification Numbers: E0, O4, R0
Keywords: Balanced Growth, Economic Growth, Productivity, Externalities, Increasing Returns, Agglomeration, Density

*Earlier versions of this paper were circulated with the title "Agglomeration and Productivity: New Estimates and Macroeconomic Implications." We thank Marco Bassetto, V.V. Chari, Jonathan Heathcote, François Ortalo-Magné, Sevi Rodriguez-Mora, José Víctor Ríos Rull, Marcelo Veracierto, and numerous seminar participants for helpful conversations and comments. The views expressed herein are those of the authors and do not necessarily represent those of the Federal Reserve Bank of Chicago or the Federal Reserve System.

1 Introduction

There is widespread agreement that cities emerge because of agglomeration effects, which can be thought of as productivity gains arising from the clustering of production and workers. Examples include superior matching of workers and jobs, knowledge spillovers that accelerate the adoption of new technologies, expanded opportunities for specialization, scale economies in the provision of common intermediate inputs, and lower transportation costs. Because agglomeration effects give rise to cities, most economic activity, and therefore growth, occurs in cities. A question that naturally arises is the extent to which local agglomeration contributes to aggregate growth. We answer this question. To do so we build and estimate a dynamic general equilibrium model of cities and growth in which local agglomeration affects per capita consumption growth. We then use panel data to estimate the key agglomeration parameters in our model. These estimates imply that agglomeration raises per capita consumption growth by about 10 percent.

Our model extends the neoclassical growth model along three dimensions. Production, consumption, and housing are location specific; local developed land is a produced, durable input into production and housing; and the economy includes local agglomeration effects as proposed by Ciccone and Hall (1996). The model features firms that produce with inputs of capital, labor, and land. Total factor productivity of firms in a location is increasing in the location's output density (output per unit of land used in production), but firms do not take this externality into account when making their decisions. We study the model's competitive equilibrium and show that along the balanced growth path, per capita consumption growth depends on the rate of increase in land prices and a parameter governing the net effect of agglomeration on productivity.

If land prices have no trend, then agglomeration has no impact on growth in our model. With growth in land prices, agglomeration affects growth via two channels. First, firms economize on land when it becomes more costly, thereby lowering growth. This is similar to the effect a declining price of equipment has on capital in Greenwood, Hercowitz, and Krusell (1997). Second, rising land prices and the ensuing curtailing of land use cause output density and therefore total factor productivity to grow faster than they would otherwise. Below we describe evidence that land prices indeed have a positive secular trend, growing 2.8 percent a year between 1980 and 2008. Combined with our estimates of the net effect of agglomeration on local productivity, this rate of increase in land prices translates into the impact of agglomeration on growth that we find.

We interpret the locations in our model as cities, broadly construed. Hence our model is related to the one analyzed by Rossi-Hansberg and Wright (2007). A key difference between our model and theirs is the role played by land. In their model land is not a factor of production. Production in a city takes place in the central business district (CBD), a point at the center of a circle. Workers demand exactly one unit of land for housing, and land is differentiated by its distance from the CBD. A city expands by adding land for workers farther away from the CBD, subject to increasing commuting costs. Growth in Rossi-Hansberg and Wright (2007) is characterized by adding land, either by expanding existing cities or adding new cities. In our model all land in a city is identical, and land is an elastic input to production and housing. There is no concept of distance because production and housing can occur anywhere in a city. Growth is characterized by adding land to existing cities. We view our model as capturing the idea that one way cities expand is by adding jobs and housing on the periphery.

Our focus on agglomeration also differs from that of Rossi-Hansberg and Wright (2007), who instead study a human capital externality. While we do not consider human capital externalities explicitly, we do account for variation in human capital in our empirical work. Following Ciccone and Peri (2006) we model labor input as a combination of low- and high-skilled workers and use this representation of labor to control for the composition of the workforce in our estimation. Doing so is important because of the clear connection between agglomeration and human capital. For example, it may be that more dense urban areas attract higher human capital workers. Not considering such behavior would lead to biases in our estimates.

Our strategy for identifying the size of agglomeration effects builds on Lucas (2001), who uses variation in land rents within a city to identify the magnitude of agglomeration effects. Similarly, we use variation across many cities. In particular, we exploit our model's predicted relationship across cities between wages, output prices, housing rents, and the composition of the workforce. Our data include Metropolitan

Statistical Area (MSA) level wage and labor input data from the Current Population Survey (CPS), output price data from the Bureau of Economic Analysis (BEA), and data on housing rental prices constructed using information from the 1990 Decennial Census of Housing (DCH) and the Bureau of Labor Statistics (BLS). This annual panel data set covers 22 MSAs from 1985 to 2006. The existence of a panel allows us to apply modern dynamic panel data techniques to estimate our model's structural parameters, in particular the Arellano and Bover (1995) forward-differencing method.

We find that local agglomeration effects are small but statistically significant. A doubling of output per unit of productive land in a city, holding all other inputs fixed, increases productivity of each firm in the city by 3 percent. However, in more dense locations, land rents are higher and firms use less land as an input to production. This lowers the net impact of agglomeration on wages and productivity to about 2 percent. While having a seemingly small effect at the local level, our model nevertheless implies that agglomeration has a substantial effect on aggregate consumption growth because of the growth in land prices.

Using a similar model of agglomeration Ciccone and Hall (1996) estimate a much larger net effect of agglomeration on local productivity. Our analysis extends theirs in several directions. First, we show how to translate estimates of the net impact of agglomeration on local productivity into the impact of agglomeration on aggregate growth. Second, we use 21 years of panel data instead of a single cross-section to estimate the size of agglomeration effects. Third, we allow for alternative uses of land at a location rather than assuming that all land at a location is used in production. Fourth, we relax the assumption that the output produced in all locations is perfectly substitutable. Below we discuss how these differences affect our results.

Ciccone and Hall (1996) can be viewed as part of a larger literature that measures the impact of agglomeration externalities on firm productivity and wages. This literature considers the effects of agglomeration on the level of output, not the growth rate as we do here. As summarized in the survey by Rosenthal and Strange (2004), researchers focus on what are called "urbanization" effects and "localization" effects. Urbanization describes the impact of a location's output or employment density on the wages and productivity of all industries in that location. Our approach to agglomeration fits into this framework. Localization describes the impact of the size of an industry in a location on wages and productivity in that industry and location. We do not consider localization, although evidence described by Henderson (2003) and Rosenthal and Strange (2003) suggests localization effects are also important determinants of local productivity.

Our estimate of a 2 percent net impact of agglomeration on productivity is below the range of estimates found in the literature. For example, Ciccone and Hall (1996) estimate this effect to be 5-6 percent, which is within the range of estimates found in recent papers, including Carlino and Voith (1992), Ciccone (2002), and Combes, Duranton, and Gobillon (2008). These papers find that a doubling of output density is associated with an increase in wages of between 3 and 8 percent. We show that estimates of the impact of agglomeration on consumption growth are quite sensitive to this number. For example, if we use the estimate of Ciccone and Hall (1996) the contribution of agglomeration to consumption growth more than doubles. Instead of taking values of the key agglomeration parameter from the literature, we estimate it ourselves because land accumulation plays a critical role in both the model and empirical analysis and because the existing literature implicitly or explicitly assumes a fixed supply of land.

The rest of the paper is organized as follows. The next section describes the production side of our model and its implied balanced growth path. Section three describes the rest of our model and verifies that the balanced growth path corresponds to a competitive equilibrium. In sections four and five we outline our strategy for estimating the key agglomeration parameter and the data underlying our estimation. Section six discusses our calibration along with estimates of the key agglomeration parameter, and section seven quantifies the impact of agglomeration on per capita consumption growth. Section eight concludes.

2 Neo-Classical Growth with Local Production

Our model extends the neo-classical growth model to include location-specific production, consumption, and housing; local developed land as a produced, durable input into production and housing; and local agglomeration effects as proposed by Ciccone and Hall (1996). We focus in this section on the production side of the economy.

This is sufficient to describe the balanced growth path of the model's competitive equilibrium.

2.1 Producers' Environment

Time is discrete and a period is equal to one year. The economy consists of a unit measure of locations, called cities. Cities are indexed on the unit interval or by their current state (s, z), where s denotes the beginning-of-period stock of developed land in the city, and z denotes exogenous city-specific productivity. Developed land is used in housing workers and for production, and it can be accumulated.[1] Cities in our model correspond to islands in Lucas and Prescott (1974). Our cities can be thought of as islands with land reclamation.

Competitive firms in each city produce an intermediate good unique to the city, y, with inputs of developed land, l_b, capital, k_b, and labor, n, using a constant-returns-to-scale production function,

$$y = a l_b^{1-\phi} k_b^{\alpha\phi} n^{(1-\alpha)\phi}, \tag{1}$$

where a is total factor productivity (TFP), $0 \leq \alpha \leq 1$, and $0 \leq \phi \leq 1$. TFP is taken by firms to be given and is specified as

$$a = z^{(1-\alpha)\phi} \left[\frac{Y}{L_b} \right]^{\frac{\lambda-1}{\lambda}},$$

where Y and L_b are total city output and developed land used in production.[2] Since we normalize the number of intermediate good producers in a city to one, in the competitive equilibrium city-specific aggregates equal those of the representative intermediate good producer in that city, $Y = y$ and $L_b = l_b$.

The ratio of output to land is called the *density of economic activity.* If $\lambda = 1$ density has no impact on productivity, and if $\lambda > 1$ firms' productivity is increasing

[1] In their study of house price dispersion, Van Nieuwerburgh and Weill (2009) consider a multi-city model with land accumulation. Land is an input into housing, but not production.

[2] If the net impact of density on production is positive, $\phi\lambda > 1$, then city output is diminishing in land. This implies that if firms were to internalize the density externality they would choose as little land as possible to produce, yielding infinite productivity. This is in contrast to Rossi-Hansberg and Wright (2007)'s model, where a human capital externality is internalized by the device of a competitive market for developers who use a scheme of subsidies and taxes to induce the first best outcome.

in density. The parameter λ measures the percent increase in productivity achieved by moving a firm holding its inputs of capital, labor, and land fixed to a city with double the output on the same amount of land. Ciccone and Hall (1996) show how this model of TFP can be derived as the reduced form of a micro-founded model.

We assume the cities share a common trend in productivity and that there are city-specific transitory shocks around the common trend. Specifically,

$$z_t = \gamma^t \tilde{z}_t, \tag{2}$$

where $\tilde{z}_t$ follows a stationary first-order Markov process, t denotes calendar time, and $\gamma \geq 1$. Productivity obeys the transition function Q, where $Q(z, z')$ is the probability that $z_{t+1} = z'$ conditional on $z_t = z$.

In each city there are also competitive housing service providers. Per capita housing services, h, are produced with developed land, l_h, and capital, k_h, according to $h = k_h^{\omega} l_h^{1-\omega}$, where $0 \leq \omega \leq 1$.

Competitive firms produce final goods by combining the output of each city using a constant-elasticity-of-substitution production function in which, for simplicity, the goods produced by each city enter symmetrically. Let $\mu_t(S, Z)$ denote the probability measure of cities with developed land in the set S and productivity shocks in the set Z, at date t. Then, final good production is given by

$$\bar{Y}_t = \left[\int y(s,z)^{\eta} \mu_t(ds, dz) \right]^{\frac{1}{\eta}}. \tag{3}$$

Normalizing the number of final goods producers to unity, $\bar{Y}_t$ corresponds to aggregate output of final goods. Final goods output is converted one-for-one into consumption, capital investment goods, and an input to new developed land, x, at the rate v.

We assume that land development costs grow at the rate $\tau \geq 1$. Land development costs increase over time if the best land is the first to be developed; the productivity of newly developed land is lower the later it is developed. If development costs are largely construction costs, then another way development costs grow is if TFP in construction grows at a slower rate than elsewhere in the economy. This seems to be the case empirically, *c.f.* Basu, Fernald, Fisher, and Kimball (2009) and Davis and Heathcote (2005).

Competitive land developers augment the stock of developed land subject to adjustment costs. The land accumulation technology is

$$s_{t+1} = \zeta(s_t, x_t), \tag{4}$$

where $\zeta(\cdot)$ is the same as the adjustment cost function proposed in Lucas and Prescott (1971) and does not affect the model's balanced growth path. In practice this requires that the derivatives of $\zeta(\cdot)$ with respect to each of its arguments are constant when evaluated along a balanced growth path. Adjustment costs introduce heterogeneity in new developed land prices into the model. This heterogeneity is necessary for there to be cross-sectional variation in land rents, which we use to estimate the key agglomeration parameter.

The developed land, housing, and labor services are traded in local markets at land rent $p^l(s,z)$, housing rent $p(s,z)$, wages $w(s,z)$, and a price for developed land $b(s,z)$. Final goods, intermediate goods, and capital services are traded in economy-wide markets at the numeraire price for consumption and capital investment goods, the price v for developed land investment goods, the intermediate good prices $q(s,z)$, and the aggregate capital rental rate r.

2.2 The Producers' Optimization Problems

We now describe the behavior of local producers of intermediate goods and housing, final goods producers, and land developers. To be clear on which variables are location-specific and which are not, we introduce dependence of variables on local state variables where appropriate.

2.2.1 Intermediate Goods and Housing Services

Profit maximization by intermediate goods producers with TFP $a(s,z)$ yields the following first-order conditions, which must be satisfied by the choices $l_b(s,z)$, $k_b(s,z)$, and $n(s,z)$,

$$p^l(s,z) = q(s,z)(1-\phi)a(s,z)l_b(s,z)^{-\phi}k_b(s,z)^{\alpha\phi}n(s,z)^{(1-\alpha)\phi}; \tag{5}$$

$$r = q(s,z)\alpha\phi a(s,z)l_b(s,z)^{1-\phi}k_b(s,z)^{\alpha\phi-1}n(s,z)^{(1-\alpha)\phi}; \tag{6}$$

$$w(s,z) = q(s,z)(1-\alpha)\phi a(s,z)l_b(s,z)^{1-\phi}k_b(s,z)^{\alpha\phi}n(s,z)^{(1-\alpha)\phi-1}. \tag{7}$$

Let $p(s,z)$ denote the price of one unit of rented housing services and $p^l(s,z)$ denote the price of a unit of rented developed land in a city with state vector (s,z). Profit maximization by housing service providers yields

$$p^l(s,z) \;=\; \omega^{\frac{\omega}{1-\omega}}(1-\omega)p(s,z)^{\frac{1}{1-\omega}}r^{\frac{-\omega}{1-\omega}} \; . \tag{8}$$

Input use by intermediate goods producers and housing service providers is limited by their supply. The local capital and developed land market-clearing conditions are

$$n(s,z)k_h(s,z) + k_b(s,z) \;=\; k(s,z), \;\; \forall\,(s,z); \tag{9}$$

$$n(s,z)l_h(s,z) + l_b(s,z) \;=\; s, \;\; \forall\,(s,z), \tag{10}$$

where $k(s,z)$ denotes the supply of capital to a city with state vector (s,z).

Market clearing in the aggregate capital and labor markets requires

$$\int k(s,z)\mu_t(ds,dz) \;\leq\; K_t; \tag{11}$$

$$\int n(s,z)\mu_t(ds,dz) \;\leq\; 1, \tag{12}$$

where K_t is the date t aggregate stock of capital and we have normalized the number of workers in the economy to unity, assuming, for simplicity, that there is no population growth.

2.2.2 Final Goods

Profit maximization of final goods producers yields

$$q(s,z) = \bar{Y}^{1-\eta}y(s,z)^{\eta-1}. \tag{13}$$

The aggregate demand for final goods is limited by their supply. Let C denote aggregate consumption and κ denote the depreciation rate of capital. The resource constraint for final goods is that total production of final goods must be at least as large as the sum of its uses:

$$C_t + K_{t+1} - (1-\kappa)K_t + v_t \int x(s,z)\mu_t(ds,dz) \leq \bar{Y}_t. \tag{14}$$

2.2.3 Land Development

Land developers in each city manage the stock of developed land to maximize profits, by renting developed land to intermediate goods producers and housing service providers and augmenting the stock of land using (4). Let $W(s, z)$ denote the expected value of a land developer in a city with state vector (s, z) and let "$'$" denote next period's value of a variable. The recursive representation of the land developer problem is then

$$W(s,z) = \max_{\{s',x\}} p^l(s,z)s - vx + \frac{1}{R}\int W(s',z')Q(z,z')dz', \tag{15}$$

subject to the land accumulation technology (4).

The variable R denotes the rate of return on a one-period bond. We do not impose that land is irreversibly developed, that is $s_{t+1} \geq s_t$, but we could easily include this constraint.

The first-order conditions for the representative land developer can be written

$$b(s,z) = \frac{1}{R}\int \left[p^l(s',z') + b(s',z')\zeta_1(s_{t+1},x_{t+1}))\right] Q(z,z')dz'; \tag{16}$$

$$v = b(s,z)\zeta_2(s_t,x_t), \tag{17}$$

where $\zeta_i(\cdot,\cdot)$ denotes the derivative of the adjustment cost function with respect to its i'th argument. Equation (16) says the price of new land in each city, $b(s, z)$, at every date, equals the discounted expected value of rent on that land, $p^l(s, z)$, plus the price of the new land in the following period. The price of the new land in the following period is its marginal product in developing new land times the price of new land in the next period. This can be thought of as the price of *existing* developed land, a concept analogous to installed capital. Equation (17) equates the marginal benefit of final goods in new land development to its cost. Free entry into land development drives $W(s, z)$ to the value of selling existing developed land to other developers.

2.3 Balanced Growth

We focus on a balanced growth path (with no uncertainty) where per capita output in all cities grows at the same rate. The variables in the model's stochastic equilibrium fluctuate around this path. This is identical to the concept of balanced growth considered by Rossi-Hansberg and Wright (2007). They show by example how this concept of balanced growth can be made consistent with empirical distributions of cities by population. Their example has i.i.d. idiosyncratic productivity shocks in a model with a locally produced, durable factor of production called human capital. We do not (need to) solve our model and so do not verify that this property is shared by our model.[3]

To derive the balanced growth path we use the market-clearing conditions, the trend growth rates of productivity and land development costs, and the local and final goods production functions. The market-clearing condition for final goods, (14), implies per capita consumption, capital, and the value of land development goods each grow at the same rate in units of the final good. Denote this common rate g_c. An implication of the final goods production function, (3), is that output growth in each city corresponds to final goods output growth. It follows that output in each city grows at the rate g_c as well. The aggregate capital market-clearing condition, (11) and the local capital market-clearing conditions, (9), imply that capital used to produce intermediate goods in each city also grows at the rate g_c. The trend in land development costs, the land accumulation technology (4), and local developed land market clearing imply that per capita developed land used in intermediate goods production grows at the rate g_c/τ in each city.

Combining these growth implications and the local production function we can derive the equilibrium value of g_c as follows. Since all cities grow at the same rate, we drop dependence of variables on local state variables. After substituting for a in equation (1), dividing through by n yields an expression involving per capita variables:

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

Divide this equation by the equivalent one for the previous period, and make appropriate substitutions for the growth rates of the variables. Solving for g_c in the resulting

[3]We can modify our model to guarantee that cities exhibit population trends relative to the aggregate population in finite samples, without changing the model's balanced growth properties. For instance, we could assume idiosyncratic productivity follows a random walk with random probability of location destruction and instantaneous replacement, as in Alvarez and Shimer (2008)'s islands model. Building in endogenous trends is more difficult, although there has been some progress in recent years. For example, the multi-sector models of Acemoglu and Guerrieri (2008) and Ngai and Pissarides (2007) might be adaptable to our purposes.

expression yields that along a balanced growth path, per capita consumption growth is given by

$$g_c = \gamma \tau^{\frac{\delta-1}{(1-\alpha)\delta}}, \tag{18}$$

where $\delta = \phi\lambda$. The parameter δ measures the effect of agglomeration net of congestion due to diminishing returns to land.

Equation (18) shows that per capita consumption growth depends on the rate of productivity growth and, if $\delta \neq 1$, the rate of growth in land development costs as well. According to equations (16) and (17) and our assumptions on $\zeta(\cdot)$, the growth rate of the relative price of new land, b, is identical to that of land development goods along a balanced growth path. If this growth rate is positive, that is, $\tau > 1$, then developed land growth will not keep up with output growth. In this case the density of economic activity grows. If agglomeration effects outweigh congestion effects, $\delta > 1$, then this mechanism provides an endogenous source of per capita consumption growth in addition to productivity growth.

Without any productivity effect of density, $\lambda = 1$ and $\delta = \phi$, equation (18) implies

$$g_c = \gamma \tau^{\frac{\phi-1}{(1-\alpha)\phi}}. \tag{19}$$

Assuming $\tau > 1$, when $0 < \phi < 1$, per capita consumption growth is lower than the rate of growth of exogenous productivity, because of decreasing returns to developed land in intermediate goods production. If agglomeration and congestion effects cancel, $\delta = 1$, then per capita consumption growth equals the rate of productivity growth, as in the simple neoclassical growth model.

Another interesting special case of our model is when the quantity of land in each location is fixed.[4] In this case the balanced growth formula is

$$g_c = \gamma^{\frac{(1-\alpha)\delta}{1-\alpha\delta}} \mu^{\frac{\delta-1}{1-\alpha\delta}}, \tag{20}$$

where μ is the rate of population growth. Population growth appears in the expression because with land fixed it drives how fast output density grows.

We use equations (18) and (19) below to quantify the contribution of agglomeration to per capita consumption growth. To do this we need to assign values to τ, ω,

[4] With fixed land the production structure is similar to that of Kiyotaki, Michaelides, and Nikolov (2008). Their model has capital, land and labor as inputs into an aggregate production function without an agglomeration externality.

α, ϕ, and δ. We use evidence on residential land prices to calibrate τ. Given a value of τ we calibrate ω, α, and ϕ using aggregate income and balance sheet data. We use panel data to estimate δ.

3 Competitive Equilibrium Growth

This section completes the description of the model. For simplicity, we initially assume zero growth in the aggregate level of productivity, the cost of land development, and the number of households and focus on a stationary equilibrium. After describing the competitive equilibrium for this case, we briefly return to the case with growth to verify that the balanced growth path we have described does in fact correspond to a competitive equilibrium. Readers interested only in our empirical findings can skip this section without much loss of continuity.

3.1 The Household's Environment

There is a representative household composed of a unit measure of identical members.[5] The large household structure allows for full risk sharing within each household, a standard device in macroeconomics for studying complete markets allocations. The household owns capital as well as producers of intermediate and final goods, and land developers. Each household member supplies one unit of labor inelastically and enjoys utility from consumption and housing.[6] The household derives income from labor and capital by allocating its members and its stock of capital across the cities. It also collects dividends from its ownership of firms. The only technology for allocating household members across cities is a random number generator. Each period the household's members are randomly allocated across cities at zero cost, after observing the current level of productivity in each city. A key assumption of the model is that household members must consume goods and housing in the same city in which they work.

[5] In the extension to heterogeneous workers described below we assume that fixed fractions of household members are either skilled or unskilled.

[6] We can include city-specific amenities and not change our results, if they enter additively in preferences.

The household takes as given the same prices as the producers in each city. The household also takes as given the law of motion for the probability distribution of developed land and productivity. The law of motion for this distribution is

$$\mu_{t+1}(S', Z') = \int_{\{(s,z):s_{t+1}(s,z)\in S'\}} Q(z, Z')\mu_t(ds, dz), \tag{21}$$

for all S' and Z'. This equation states that the total number of cities at date $t+1$ with developed land in the set S' and productivity in the set Z' is given by the total of all cities that transit from their current productivity to productivity in Z' and produce developed land so that $s' = s_{t+1}(s, z)$ is in S'.

3.2 The Household's Optimization Problem

The household maximizes the expected present value of average household member utility, where household member utility in any given period is logarithmically separable in consumption and housing. Separability implies the household perfectly insures itself against consumption risk. Therefore every household member receives the same level of consumption, C, which corresponds to aggregate consumption when we normalize the number of household members and households to unity. Let $V(K, \mu)$ denote the expected utility of the household with capital K and distribution of land and productivity μ. The recursive representation of the household's problem is

$$V(K) = \tag{22}$$

$$\max_{\left\{\substack{C,K',h(s,z)\\ n(s,z)}\right\}} \left\{ \int n(s,z)\left[\ln C + \psi \ln h(s,z)\right]\mu(ds,dz) + \beta \int V(K')\mu'(ds',dz') \right\}$$

subject to

$$\begin{aligned} C + K' - (1-\kappa)K \quad &+ \int p(s,z)n(s,z)h(s,z)\mu(ds,dz) \\ &\leq \; rK + \int w(s,z)n(s,z)\mu(ds,dz) + \pi, \end{aligned} \tag{23}$$

the employment constraint (12) and the transition equation (21), where $\psi \geq 0$.[7] The left-hand side of the household's budget constraint includes the household's expenditures on consumption, new capital, and housing. The right-hand side includes income

[7] To conserve on notation we have not distinguished between the demand and supply of final goods or factors of production.

from capital, labor, and dividends from ownership of intermediate goods producers, housing service providers, land developers, and final goods producers, π.

The first-order conditions for K', $h(s,z)$ and $n(s,z)$ are

$$1 = \beta \frac{C}{C'} \left[r' + 1 - \kappa \right]; \tag{24}$$

$$\frac{\psi}{h(s,z)} = \frac{p(s,z)}{C}, \ \forall (s,z); \tag{25}$$

$$\ln C + \psi \ln h(s,z) = \theta + \frac{1}{C} \left[p(s,z) h(s,z) - w(s,z) \right], \ \forall (s,z), \tag{26}$$

where θ is the Lagrange multiplier for (12). The capital accumulation condition (24) is familiar from the simple neoclassical growth model. The simple form of equation (25) is due to our specification of preferences. Its implication of a constant housing share of consumption is supported by evidence in Davis and Ortalo-Magné (2008). Equation (26) is unique to our model. Combined, equations (25) and (26) imply that utility is not identical for each household member but rather household members assigned to relatively high-wage cities enjoy a relatively low quantity of housing and thus relatively low utility. This feature of our model contrasts with the static Roback (1982) model where *ex post* utility of workers is equated across cities. In our dynamic model with complete markets *ex ante* expected utility is equalized.

We conclude this subsection by noting that equation (26) reveals the important role housing plays in our model. Consider the case where housing is not valued by household members, $\psi = 0$. Equation (26) is satisfied in this case only if wages are equalized across locations. Since our estimation procedure relies on cross-sectional variation in wages, we require that housing is valued, $\psi > 0$.

3.3 Stationary Competitive Equilibrium

A *stationary competitive equilibrium* consists of value functions $V(K)$ and $W(s,z)$; decision rules for consumption $g_C(K)$, capital accumulation $g_K(K)$, land development $s'(s,z)$, housing $h(s,z)$, land in production $l(s,z)$, capital in production $k(s,z)$, and workers $n(s,z)$; prices $\{q(s,z), p^l(s,z), w(s,z), p(s,z), b(s,z), r, v\}$; aggregate quantities $\{C, K'\}$; and a measure $\mu(s,z)$, such that

1. Given prices, households maximize expected utility, so that $V(K,\mu)$ solves the

household problem as given by (22) and equations (23)–(26) are satisfied;

2. Given prices, producers maximize profits, so that $W(s,z)$ solves (15) and equations (3)–(7), (13), (16), and (17) are satisfied;

3. The discount factor $1/R$ is given by β;

4. Aggregates are consistent with individual behavior, so that $\mu(s,z)$ is generated by (21);

5. Markets clear so that equations (9)–(12) and (14) are satisfied.

When $\lambda = 1$ the competitive equilibrium corresponds to the solution of a concave programming problem. Standard arguments can be used to prove in this case that there exists a unique competitive equilibrium and the allocations are Pareto-efficient.

We have not found conditions under which an equilibrium exists with density effects, $\lambda > 1$. However, using the argument in Kehoe, Levine, and Romer (1992) we know that for λ sufficiently close to unity, there exists a unique equilibrium, because the model without density effects has a unique equilibrium. Equilibrium existence and uniqueness with the size of density effects we estimate remain open questions. We do not need to find a solution of the model to estimate the impact of density on local productivity or quantify the effect of agglomeration on per capita consumption growth, and so do not do so in this paper.

So far in this section we have assumed no growth. This is without loss of generality because the preceding discussion applies to a version of the model specified in terms of variables scaled so that they are stationary along a balanced growth path: the balanced growth path corresponds to a competitive equilibrium. This is established by verifying that the conditions stated in the definition of the competitive equilibrium are compatible with the balanced growth path. Since we derived this path using the market-clearing conditions, these conditions are obviously compatible. It remains to establish the compatibility of the first-order conditions of the firms' and household problems. This is accomplished by verifying that if the first-order conditions are satisfied by the scaled variables, they are also satisfied by the unscaled variables.[8]

[8] This is straightforward for all but the first-order condition for allocating household members to cities, equation (26). Since C and h appear in both levels and logs in this equation, it must be

4 Empirical Strategy

In this section, we describe our procedure to estimate the key agglomeration parameter δ given estimates for α and ω. We discuss the other parameters we need to quantify the impact of agglomeration on consumption growth below in sections 6.1. As in Ciccone and Hall (1996), we derive the equation underlying our estimation of δ using production efficiency alone. Consequently, our estimates do not rely on the additional assumptions, such as preferences and insurance arrangements, we need to specify our model completely. We begin by showing how the intermediate goods producers' efficiency conditions can be used to identify δ. Since the composition of the workforce is important in our estimation, we then extend our model to include heterogeneous workers and show how this affects how we identify δ. We then derive the basic form of our estimation equation. Finally, we describe how our estimation procedure addresses endogeneity and fixed effects.

4.1 Identifying δ

Consider any two cities, indexed i and j, in some arbitrary period. Solving for output in (1) and dividing output in city i by that in city j yields

$$\frac{y_i}{y_j} = \left[\frac{z_i}{z_j}\right]^{\delta(1-\alpha)} \left[\frac{l_i}{l_j}\right]^{1-\delta} \left[\frac{k_i}{k_j}\right]^{\delta\alpha} \left[\frac{n_i}{n_j}\right]^{\delta(1-\alpha)}.$$

Using (6) to eliminate the capital stock term in this last equation we have

$$\frac{y_i/l_i}{y_j/l_j} = \left[\frac{z_i}{z_j}\right]^{\frac{\delta(1-\alpha)}{1-\delta\alpha}} \left[\frac{n_i/l_i}{n_j/l_j}\right]^{\frac{\delta(1-\alpha)}{1-\delta\alpha}} \left[\frac{q_i}{q_j}\right]^{\frac{\delta\alpha}{1-\delta\alpha}}. \tag{27}$$

Equation (27) is closely related to equation (19) in Ciccone and Hall (1996), which is the basis for their estimation equation. It differs in two key respects. First, because Ciccone and Hall (1996) assume intermediate goods are perfect substitutes

handled differently. Notice that this equation is the only one in which the Lagrange multiplier on the allocation of household members across cities, θ, appears. In other words, equation (26) determines the equilibrium value of θ, assuming an equilibrium exists. Equation (26) implies that $\theta > 0$ at all dates along a balanced growth path, as long as labor income always exceeds the value of housing services in each city. This condition will be satisfied in any plausible calibration. We conclude that equation (26) is compatible with balanced growth for plausible calibrations.

in producing final goods, $\eta = 1$, their equation does not contain any output prices. So if $\eta < 1$ their equation is subject to an omitted variable bias. Second, they assume all land in a given location is used in production while we have competing uses for land. Since the allocation of land between residential and non-residential uses in U.S. cities is unavailable, we cannot use (27) as a basis for estimation. Further complicating our analysis, annual data on real city output that is produced by the BEA is available for too short a time (2001–2006) to be useful to us.

We eliminate employment, developed land, and output from (27) using (5) and (7). Substituting for developed land rent using (8) yields the equation underlying our estimation of δ :

$$\frac{w_i}{w_j} = \frac{z_i}{z_j} \left[\frac{p_i}{p_j}\right]^{\frac{1}{1-\omega}\frac{\delta-1}{\delta(1-\alpha)}} \left[\frac{q_i}{q_j}\right]^{\frac{1}{\delta(1-\alpha)}}. \tag{28}$$

Notice that the exponent on housing rents is proportional to the exponent on the growth rate of land prices in the balanced growth formula, equation (18). The factor of proportionality involves the share of capital in housing, ω. So, with an estimate of ω in hand, we could, in principle, quantify the growth implications of agglomeration without separately estimating δ.

4.2 Extension to Heterogeneous Workers

So far we have assumed that workers are homogeneous, which is a questionable assumption from an empirical perspective. As emphasized in the introduction, not accounting for cross-sectional variation in the distribution of human capital could lead to biased estimates of density effects. To address this issue we follow Ciccone and Peri (2006). They consider "skilled" and "unskilled" workers as imperfect substitutes in producing total labor services. We now derive a version of the equation underlying our estimation that incorporates heterogeneous workers. The way we incorporate heterogeneous workers does not affect the balanced growth path of our model.

Suppose that the effective labor input in (1) is a constant elasticity of substitution composite of unskilled, n_u, and skilled labor, n_e:

$$n = \left[\sigma n_u^{\xi} + (1-\sigma) n_e^{\xi}\right]^{1/\xi},$$

where $0 < \sigma < 1$ and $\xi \leq 1$. Composite labor satisfies (7) as before. Let w_u and w_e denote the wages of unskilled and skilled workers. The first-order conditions for intermediate goods producers' choices of unskilled and skilled labor can be used to express the wages of skilled workers as

$$w_e = (1-\sigma)\sigma^{1/\xi-1}(1-\alpha)\phi w \chi^{1/\xi-1} m^{\xi-1},$$

where w denotes the implicit wage for the composite labor input, n, $\chi \equiv (w_u n_u + w_e n_e)/(w_u n_u)$, and $m \equiv n_e/n_u$. Substituting for composite wages using (28), we have

$$\frac{w_{ei}}{w_{ej}} = \frac{z_i}{z_j}\left[\frac{p_i}{p_j}\right]^{\frac{1}{1-\omega}\frac{\delta-1}{\delta(1-\alpha)}}\left[\frac{q_i}{q_j}\right]^{\frac{1}{\delta(1-\alpha)}}\left[\frac{\chi_i}{\chi_j}\right]^{1/\xi-1}\left[\frac{m_i}{m_j}\right]^{\xi-1}, \tag{29}$$

for any two cities i and j. Equation (29) reduces to equation (28) if $\xi = 1$, that is, if unskilled and skilled labor are perfect substitutes. The variables χ and m are relatively straightforward to measure so this equation can be used to estimate δ taking into account worker heterogeneity.

4.3 The Estimation Equation

Take logs of equation (29) and rearrange terms:

$$\begin{aligned}\ln(w_{ei}) - \ln(w_{ej}) \;=\; & \frac{1}{1-\omega}\frac{\delta-1}{\delta(1-\alpha)}\left[\ln(p_i) - \ln(p_j)\right] + \frac{1}{\delta(1-\alpha)}\left[\ln(q_i) - \ln(q_j)\right] \\ & + \frac{1-\xi}{\xi}\left[\ln(\chi_i) - \ln(\chi_j)\right] + (\xi - 1)\left[\ln(m_i) - \ln(m_j)\right] \\ & + \ln(z_i) - \ln(z_j)\end{aligned}$$

This equation holds for any two cities i and j at all dates. Therefore, it must also hold for city i and the average of a sample of size N of cities. Define a "hat" over a variable for city i to mean the difference of the log of that variable and the average of the same logged variable for all N cities in the sample, for example,

$$\hat{w}_{ei} \equiv \ln(w_{ei}) - \frac{1}{N}\sum_{j=1}^{N}\ln(w_{ej}).$$

Notice that by construction hatted variables are mean zero in every year. Re-introducing the time subscript, it follows that

$$\hat{w}_{eit} = \frac{1}{1-\omega}\frac{\delta-1}{\delta(1-\alpha)}\hat{p}_{it} + \frac{1}{\delta(1-\alpha)}\hat{q}_{it} + \frac{1-\xi}{\xi}\hat{\chi}_{it} + (\xi-1)\,\hat{m}_{it-1} + \hat{z}_{it}, \tag{30}$$

which also holds at all dates. This is our baseline estimation equation. Because all variables are expressed as deviations from year averages, our identification of the model's parameters is based on changes in the cross-sectional distribution of our panel that occur over time.

We have data on all variables in this equation except for $\hat{z}_{it}$. If we were unconcerned about the correlation of $\hat{z}_{it}$ with the other right-hand-side variables, and were not interested in the magnitudes of the structural parameters, then we could obtain a consistent estimate of the coefficient on housing rents by ordinary least squares (OLS). With an estimate of ω in hand we could then quantify the impact of agglomeration on growth. However, the model predicts that $\hat{z}_{it}$ and the other right-hand-side variables are correlated and our theory imposes restrictions on the coefficients that should be satisfied for our findings to be credible. So, to estimate the structural parameters we need to use a non-linear instrumental variables approach. If we had only one wave of data, we would have to argue for an instrument that is correlated with the observable variables and uncorrelated with the $\hat{z}_{it}$ term and then we could apply non-linear least squares to obtain consistent estimates. This is the empirical approach adopted by Ciccone and Hall (1996). Since we have panel data with a long time dimension, we use a different estimation strategy.

4.4 Addressing Endogeneity

We address endogeneity by making a parametric assumption on the evolution of city-specific productivity. The stationary component in (2) is assumed to be the sum of two random variables, one that is a city-specific first-order auto-regressive process, and one that is common to all cities. We include this second term to make it clear that our empirical procedure is consistent with the presence of an aggregate productivity shock. For every city i and all $t \geq 1$, $\ln(z_{it})$ evolves according to

$$\begin{aligned} \ln(z_{it}) &= \ln(z_{i0}) + \gamma t + \ln(\tilde{z}_{it}) + u_t; \\ \ln(\tilde{z}_{it}) &= \rho \ln(\tilde{z}_{it-1}) + e_{it}, \end{aligned}$$

where γ is the deterministic rate of growth from (2), u_t is the economy-wide stochastic component of productivity, $-1 < \rho < 1$ is the autocorrelation coefficient for city-specific productivity, and e_{it} is a city-specific i.i.d. shock that is orthogonal to all variables dated $t-1$ and earlier. Both γ and ρ are common to all cities. The variable z_{i0} is the initial level of productivity specific to city i.

We now use our assumptions on the evolution of $\tilde{z}_{it}$ to manipulate (30) into an equation with an error term that is uncorrelated with all variables from previous years. This forms the basis of our instrumental variables estimation. Notice that our assumptions on the evolution of z_{it} and $\tilde{z}_{it}$ imply that the growth term and the economy-wide stochastic term drop out of the deviation of $\ln(z_{it})$ from its cross-section average, $\hat{z}_{it}$. Furthermore,

$$\hat{z}_{it} - \rho\hat{z}_{it-1} \quad = \quad (1-\rho)\,\hat{z}_{i0} + \epsilon_{it}, \tag{31}$$

where $\epsilon_{it} \equiv [e_{it} - \bar{e}_t]$ with $\bar{e}_t$ equal to the period t average of city-specific productivity shocks for the N cities in our sample. Now, from each variable in equation (30) subtract ρ times its once-lagged value. This is a valid operation since equation (30) holds at all dates. Then, using equation (31), we have

$$\begin{aligned}\hat{w}_{eit} \quad = \quad & (1-\rho)\hat{z}_{i0} + \rho\hat{w}_{eit-1} + \frac{1}{1-\omega}\frac{\delta-1}{\delta(1-\alpha)}\left[\hat{p}_{it} - \rho\hat{p}_{it-1}\right] + \frac{1}{\delta(1-\alpha)}\left[\hat{q}_{it} - \rho\hat{q}_{it-1}\right] \\ & + \frac{1-\xi}{\xi}\left[\hat{\chi}_{it} - \rho\hat{\chi}_{it-1}\right] + (\xi-1)\left[\hat{m}_{it} - \rho\hat{m}_{it-1}\right] + \epsilon_{it}. \end{aligned} \tag{32}$$

To estimate equation (32) we need to find a set of instruments that are correlated with the wage, price, and skill variables and are uncorrelated with ϵ_{it}. Finding valid instruments is straightforward because ϵ_{it} is an i.i.d. shock; any variable dated $t-1$ or earlier is potentially a valid instrument. In addition, we need to address the unobserved variable $\hat{z}_{i0}$. This variable is a "fixed effect" in the sense that it varies across cities but is fixed over time in each city.

4.5 Addressing Fixed Effects

Several strategies have been proposed that can in principle address the city fixed effects. The most common one involves eliminating the fixed effect by taking first differences. A second strategy is to assume that the level of the fixed effect, $\hat{z}_{i0}$, is uncorrelated with the first-differences of all model variables. The application of this method to our case would involve using equation (32) with lagged growth rates of model variables as instruments. A third widely used strategy combines the first two strategies. This approach was originally proposed by Blundell and Bond (2000). We are uncomfortable assuming that the fixed effect is uncorrelated with lagged changes of model variables, since we have no theory suggesting this to be the case. Therefore we do not use the Blundell and Bond (2000) estimation procedure.

When more waves of data are available than are the minimum necessary to implement the first-difference strategy, taking first differences does not use all the available information on the fixed effect and so is inefficient. The procedure proposed by Arellano and Bover (1995) takes advantage of the additional information on the fixed effects that comes with long panel data and so this is what we use. In this approach, each time-t variable in equation (32) is expressed as a deviation from the average of all future observations for city i in the sample. For any variable x_t with observations $t = 1, \dots, T$, define the Arellano-Bover difference operator as of date t, Δ_t, as follows:

$$\Delta_t x_t = \varphi_t \left[x_t - \frac{1}{T-t} \sum_{s>t} x_s \right],$$

where

$$\varphi_t = \left(\frac{T-t}{T-t+1} \right)^{1/2}.$$

Applying the Arellano-Bover difference operator to (32) yields

$$\begin{aligned} \Delta_t \hat{w}_{eit} &= \rho \Delta_t \hat{w}_{eit-1} + \frac{1}{1-\omega} \frac{\delta - 1}{\delta(1-\alpha)} \left[\Delta_t \hat{p}_{it} - \rho \Delta_t \hat{p}_{it-1} \right] \\ &\quad + \frac{1}{\delta(1-\alpha)} \left[\Delta_t \hat{q}_{it} - \rho \Delta_t \hat{q}_{it-1} \right] + \frac{1-\xi}{\xi} \left[\Delta_t \hat{\chi}_{it} - \rho \Delta_t \hat{\chi}_{it-1} \right] \\ &\quad + (\xi - 1) \left[\Delta_t m_{it} - \rho \Delta_t \hat{m}_{it-1} \right] + \Delta_t \epsilon_{it}. \end{aligned} \tag{33}$$

Including the weight term φ_t in Δ_t guarantees that the error term $\Delta_t \epsilon_{it}$ in (33) has a constant variance.

We use generalized method of moments (GMM) to estimate equation (33) using the levels of $\hat{w}_{eit-s}$, $\hat{p}_{it-s}$, $\hat{q}_{it-s}$, $\hat{\chi}_{it-s}$, and $\hat{m}_{it-s}$ for $s \in \{3, 4\}$ as instruments. Lags $s \geq 5$ are not used as instruments because typically they do not add much information. We do not use lags $s \in \{1, 2\}$ to accommodate classical measurement error in all the variables. Monte Carlo studies by Blundell and Bond (2000) and Windmeijer

(2005) suggest that procedures that use lagged levels of variables as instruments on differenced data may yield biased estimates of parameters and standard errors. To understand if these results apply to our data, we conduct a Monte Carlo study based on our data. We discuss the implications of this study when describing our empirical results. The details and results of the Monte Carlo study are outlined in the appendix.

5 Data for Estimating δ

To implement our GMM estimation strategy we need city-level data on housing rents, output prices, and wages and hours of unskilled and skilled workers. A city in our data is the Census Bureau's 1990 definition of an MSA. All MSAs are defined in terms of counties, so the boundaries of the MSAs in our sample are fixed.[9] While the boundaries are fixed, that does not mean the quantity of developed land is fixed. Indeed, the MSAs in our sample all contain land that is initially undeveloped (farm and recreational land) and eventually is developed for urban purposes. Therefore, the new land development mechanism present in our model also should be present in our data. This section summarizes our MSA-level data and conducts some preliminary analysis of it. A more detailed description of our data is contained in the appendix.

5.1 Housing, Labor, and Output Variables

We construct annual data on housing rents by combining micro-level data from the 1990 Decennial Census of Housing (DCH) with housing rental price indexes from the Bureau of Labor Statistics (BLS). Specifically, we use the data from the 1990 DCH to estimate the level of housing rents by MSA in 1990, and then use MSA-specific rental price indexes from the BLS to extrapolate the rental price level backward to 1985 and forward to 2006.

We construct our wage and employment variables from the March Current Population Survey (CPS). Skilled workers are identified as those workers with at least four years of college. Unskilled workers are identified as workers with less than four years

[9] The MSA may not be the right geography corresponding to our model. Rozenfeld, Rybski, Gabaix, and Makse (2009) propose an alternative geography based on satellite imagery. We cannot use their geography because our data are organized by MSA.

of college. We measure wages for a given type of worker as total wages divided by total hours for that type of worker.

We use data from the Bureau of Economic Analysis (BEA) to construct price indexes for output by MSA. These indexes are created as weighted averages of industry-specific price indexes where the weights are based on the share of total earnings paid to employees by each industry in an MSA. Because the mix of industries varies by MSA, and price indexes vary by industry, our price index for output also varies by MSA. We normalize the price index for MSA-level output to 1.0 in 1969 in every MSA. This arbitrary normalization introduces another MSA-level fixed effect that is accounted for by the Arellano and Bover (1995) procedure.

We drop MSAs with incomplete or missing wage, employment, output price, or housing rent data. Largely because the BLS tracks rent indexes for only 27 MSAs, this leaves us with data covering the 1985–2006 period for 22 MSAs. The 22 MSAs are, roughly speaking, the 22 largest MSAs by population in the US. These MSAs include 36 percent of the total US population.

Since all our econometric specifications are based on within-year deviations from averages, we do not need to adjust any of our price variables for overall price inflation. Consequently, we use nominal wage, housing rent, and output price variables directly in our analysis. This is a valid procedure as long as the variation in consumption goods prices across cities is captured by the fixed effect.

5.2 Summary Statistics

Table 1 shows standard deviations and correlations of our data. By construction, all variables have zero mean in each year and thus are zero mean unconditionally. The reported standard deviations and correlations are estimated from the complete panel; they are not averages of annual statistics.[10] There are two features of this table worth noting.

First, the standard deviation of housing rents, $\hat{p}_{it}$, is larger than the standard

[10]The standard deviation of output prices, $\hat{q}_{it}$, cannot be interpreted literally, as output prices in every MSA are arbitrarily normalized to equal 1.0 in 1969. We could have increased the standard deviation by setting the normalized price level to vary across MSAs.

deviation of high-skill wages $\hat{w}_{eit}$. This is predicted by our model. To see why, notice that equations (25) and (26) imply that that the percentage deviation of housing rents between any two metro areas i and j approximately satisfies

$$\frac{p_{it} - p_{jt}}{p_{jt}} = \left(\frac{w_{jt}}{C_t}\right)\left(\frac{1}{\psi}\right)\left(\frac{w_{it} - w_{jt}}{w_{jt}}\right) ,$$

where we approximate $\ln(p_{it}/p_{jt})$ as $(p_{it} - p_{jt})/p_{jt}$. As long as $w_{jt}/c_t > \psi$ then the percentage deviation of rents is greater than the percentage deviation of wages. Davis and Ortalo-Magné (2008) report that the average expenditure share on housing in the United States is 24 percent. According to (25), this suggests $\psi = 0.24$. Since wages everywhere in our sample are greater than 24 percent of the level of consumption, housing rents must be more dispersed than wages paid to labor. In our extension of the model to heterogeneous workers, the same argument can be used to show that housing rents must be more dispersed than skilled wages as long as the ratio of skilled wages to consumption is greater than the household expenditure share on housing.

Second, skilled wages, $\hat{w}_{eit}$ are positively correlated with housing rents, p_{it}, and the ratio of skilled to unskill workers, m_{it}. It is straightforward to show using the first-order and market-clearing conditions that these features of the data are also predicted by the model.

5.3 Preliminary Regression Analysis

Table 2 shows the coefficient estimates and robust standard errors of simple OLS regressions of wages of high-skill workers on housing rents, $\hat{p}_{it}$, output prices, $\hat{q}_{it}$, and the two variables related to the skill composition of the workforce, $\hat{\chi}_{it}$ and $\hat{m}_{it}$. The first column corresponds to a regression in levels including MSA-specific fixed effects; and the second column corresponds to a regression in differences without fixed effects.

From these regressions, we observe three features of the data. First, the model variables account for a significant fraction of the variation in skilled wages. In the regressions in levels the R^2 statistic is 0.79, compared to 0.62 when the regression includes only fixed effects. In the difference regression, where fixed effects have been removed, the R^2 statistic is 0.37.

Second, the coefficient on housing rents is estimated to be positive in both regressions. In the levels regression it is also statistically significant. Since the coefficient on housing rents is proportional to the exponent determining whether there is a role for agglomeration in aggregate growth, these findings suggest we should find a significant role when we examine our structural estimates.

Third, the coefficients on output prices, $\hat{q}_{it}$, are imprecisely estimated and vary substantially across the two regressions. This suggests output prices (as we measure them) do not account for much of the variation in high-skill wages after controlling for skill composition and housing rents. However, we show below that including output prices has a noticeable impact on our estimates of the labor substitutability and serial correlation parameters, ξ and ρ.

6 Estimates of the Model's Parameters

This section describes our estimates of the model's parameters. The first subsection describes how we calibrate the growth rate of the developed land technology, τ, capital's share of non-land income, α, capital's share of housing, ω, and capital plus labor's share of income, ϕ. We then describe our baseline estimates of δ, and after this we show how two key model assumptions influence these estimates.

6.1 Calibrated Parameters

We begin by considering τ. This parameter is an input to calibrating $\{\alpha, \omega, \phi\}$ and is important in our balanced growth formula. Along our model's balanced growth path the consumption price of an efficiency unit of developed land grows at the rate τ. The effect of agglomeration on growth is increasing in τ, so it is useful to have an upper bound on its magnitude. Notice that since the share of income going to land is constant along a balanced growth path, the quantity of developed land grows at the rate $g_c\mu/\tau$, for consumption growth g_c and rate of population growth μ. Since the observed stock of developed land is growing, this relationship puts an upper bound on τ equal to $g_c\mu$.[11] Based on the values for g_c and μ discussed below, this upper

[11]The US Department of Agriculture uses satellite imagery to estimate the developed land area of the US, defined as "large urban and built-up areas, small built-up areas, and rural transportation

bound is 3.1 percent.

The ideal way to measure τ is with direct observations on the growth of developed land prices. We are aware of only two publicly available series on land prices, the price of farmland per acre from the USDA and estimates of residential land prices by Davis and Heathcote (2007). Figure 1 displays these data in logs for the years 1950–2008. We convert nominal prices into real prices using the deflator corresponding to our measure of consumption discussed below.

Farmland is not the kind of land we have in mind in our model, although farm land is a common source of new land to develop. We view growth of farmland prices as providing a lower bound on τ. We do not model farmland prices. Still, if agglomeration makes cities grow faster than areas without agglomeration, then there is a tendency for land prices in cities to grow faster than farm land prices as long as exogenous farm TFP does not grow faster than in the rest of the economy. Average growth of farmland prices from 1951 to 2008 is 2.6 (robust standard error 1.2) percent per year.

Residential land data is more suitable for our purposes since it is developed land. While the farmland prices are based on direct observations, residential land prices need to be inferred from house prices. This is because land in residential use is typically bundled with a structure. Average growth of residential land prices from 1951 to 2008 is 3.6 (0.2) percent. This is larger than our upper bound. For reasons discussed below, we think a sample beginning in 1980 is appropriate when considering residential land prices. Over the 1980 to 2008 sample the average growth rate of residential land prices is 2.8 (0.2) percent, within our bounds. Consequently we calibrate $\tau = 1.028$.

The parameters α and ϕ are calibrated as follows. First, we estimate the non-labor share of aggregate income, which corresponds to $1 - \phi + \alpha\phi$ in our model. Next, we estimate the share of non-labor income attributable to land, which in our model is the ratio of $1 - \phi$ to $1 - \phi + \alpha\phi$. Finally, we use the estimates of these two shares to solve for the two unknowns, α and ϕ.

The non-labor share of income is measured as follows. Following Cooley and

land." Data from the USDA's 2003 National Resources Inventory has developed land in the US growing by 1.88 percent per year between 1982 and 2003.

Prescott (1995), we assume that labor's share of "ambiguous income" is the same as that for "unambiguous income." To calculate unambiguous income we subtract from aggregate income those components of income where the factor payments are ambiguous, including taxes less subsidies, proprietors' income, and the surplus of government enterprises. We use the ratio of compensation of employees to unambiguous income to measure labor's income share and one minus labor's income share to measure the non-labor income share.[12] Over the 1959–2005 period for which data are available, these calculations yield a non-labor income share of about 0.31.

We calibrate land's share of non-labor income using our model and an estimate of the ratio of the value of land to the value of all tangible assets in the nonfinancial corporate sector from the Flow of Funds Accounts. Tangible assets include land, structures, equipment and software, and inventories.[13] Over the 1952–2009 period for which data are available, the average ratio of land value to tangible assets value is about 15 percent. We convert this value ratio into an income share using the intertemporal first-order conditions for land and capital. Along a balanced growth path, equation (16) implies the value of an efficiency unit of land equals $p^l/(R-\tau)$ and equation (24) implies the value of a unit of capital equals $r/(R+\kappa-1)$. It is straightforward to manipulate these two expressions to derive the desired income ratio. For these calculations we require values of τ, R, and κ. We use our calibrated τ, assume $R = 1.05$, and for κ we use the depreciation rate on business capital from Davis and Heathcote (2005), $\kappa = 0.0557$. The resulting estimate of land's share of non-labor income is 3.5 percent. Along with the non-labor share calculated above this yields $\{\alpha, \phi\} = \{0.3, 0.99\}$. The estimate of ϕ is in the range reported by Ciccone (2002).

We calibrate land's share of housing rents, $1-\omega$, using the same strategy we just used to calibrate land's share of non-labor income. Davis and Heathcote (2007) estimate a time series for the ratio of land value to the market value of housing over the period 1950–2008. Over the entire sample the ratio's average value is 0.27. However, over the period 1950–1979 the ratio displays a clear upward trend. From 1980–

[12]These numbers are obtained from NIPA Table 1.10 (before the 2009 revision.) The ambiguous income components are lines 9, 10, 15 and 22 of this table and labor income is line 2.

[13]These numbers are obtained from Table B.102. The market value of land is the market value of real estate (line 3) less the replacement cost value of structures used for nonresidential purposes (line 34). The value of all tangible assets held by the nonfinancial corporate sector is line 2.

2008, despite large variations, the series appears roughly stable around an average value of 0.35. Since our results are sensitive to ω we consider two values for this parameter, based on these two estimates of the ratio of land's share of housing rents. Using the intertemporal first-order conditions as above, these estimates yield land's share of housing equal to 0.15 and 0.11, corresponding to the ratios 0.35 and 0.27, so we calibrate $\omega = 0.85$ and $\omega = 0.89$.[14] The larger value for ω equals the Census Bureau estimate of the share of new home value attributable to the value of structures described in Davis and Heathcote (2005).

We conclude this subsection by using aggregate CPI data on housing rents to validate our calibrated values for τ and ω. According to equation (8) the growth rate of housing rents equals $(1-\omega)$ multiplied by the growth rate of land rents. Since land rents grow at the same rate as land values along our model's balanced growth path, it follows that $(1-\omega)\tau$ evaluated at our calibrated values should correspond closely to trend growth in measured housing rents. With $\tau = 1.028$, the predicted growth rates for housing rents is highest at 0.42 percent a year at $\omega = 0.85$. Using the same deflator underlying our measure of real residential land prices, trend growth in real CPI housing rents between 1981 and 2008 is 0.82 percent. When we use the CPI deflator for urban consumers (excluding housing services) to deflate CPI rents, trend growth in real housing rents is 0.61 percent a year. If the quality adjustment for housing rents is worse than for other goods and services in the CPI, then these growth rates are upward biased. Given the underlying uncertainty involved, we view these empirical growth rates for housing rents as validating our calibration with $\omega = 0.85$.

6.2 GMM Estimates

Table 3 reports parameter estimates, asymptotic standard errors, and specification tests when we estimate $\{\delta, \xi, \rho\}$ by GMM using the procedure of Arellano and Bover (1995). These estimates are based on $\{\alpha, \phi\} = \{0.3, 0.99\}$. Given the previous discussion, we regard $\omega = 0.85$ as the most plausible value for this parameter. The estimates for this baseline case are in the first column. The second column reports

[14] We use the same inputs to this calculation as for land's share of non-labor income except that we allow for a different depreciation rate for housing capital compared to business capital. We use $\kappa = 0.0157$, the value reported in Davis and Heathcote (2005). It is straightforward to extend our model to allow for two different kinds of capital.

how our estimates change when $\omega = 0.89$.

We employ three specification tests, summarized in the bottom half of Table 3. The first is the J-test of the over-identifying restrictions due to Hansen (1982) and Sargan (1958). The p-values corresponding to these test statistics indicate that the model is never rejected by the J-test. Arellano and Bond (1991) argue that, in our context, the power of the J-test to detect model misspecification can be quite low. As a more powerful alternative they suggest a test of the serial correlation of the residuals. If the model is correctly specified, the residuals of our estimating equation should exhibit autocorrelation only up to order one. For each case we report the $m2$ test of second-order serial correlation from Arellano and Bond (1991), and the analogous test for third-order serial correlation, denoted $m3$. The p-values corresponding to these test statistics indicate that the null hypotheses that the residuals do not display second- or third-order serial correlation are never rejected.

The point estimates of δ, 1.020 and 1.015, corresponding to ω equal to 0.85 and 0.89, both exceed unity. We focus on our baseline estimate, $\delta = 1.020$. Our Monte Carlo study, described in the appendix, suggests this estimate is essentially unbiased. Recall that δ is the product of ϕ, capital's plus labor's share of production, and λ, the density parameter. The parameter δ is thus the impact of the density externality on production, net of the fact that land is also an input in production. Our estimate $\delta > 1$ implies that a reduction in the amount of land used in production, on-net, increases output; the positive impact from the change in output density more than compensates for the reduction of one of the inputs in production. A value of δ greater than unity also implies that agglomeration increases aggregate growth beyond that implied by exogenous technological change alone. The asymptotic standard error of the baseline estimate δ is 0.002. Our Monte Carlo study suggests that the reported standard error on δ may be too low by a factor of about 2.4, such that the "true" standard error of our baseline estimate is closer to 0.005. If this is the case, we easily reject that $\delta = 1$. Similar calculations imply that our second estimate of δ is also significantly greater than unity.

With an estimate of ϕ, we can determine the direct impact of density on productivity as $\lambda = \delta/\phi$. Given $\phi = 0.99$ we easily reject the hypothesis that $\delta = \phi$ so that agglomeration increases productivity in a statistically significant way. Our baseline estimate of δ and $\phi = 0.99$ imply $\lambda = 1.03$. A firm located in an area with twice the output density will produce $(\lambda - 1)/\lambda = 2.9$ percent more output than an otherwise identical firm, holding all other inputs to production constant.

We estimate ξ and ρ in all cases to be about 0.56 and 0.57. Our Monte Carlo study suggests that if the true value of ρ is 0.6, then both of these estimates are essentially unbiased. This study also suggests that our reported standard errors are too low by a factors of 2 for ξ and 1.4 for ρ.

Our estimate of $\xi = 0.56$ implies an elasticity of substitution of 2.3 between unskilled and skilled labor types. This is higher than the range of recent estimates of the elasticity of substitution between college- and non-college educated workers of 1.3 to 1.7 as reported by Autor, Katz, and Krueger (1998).[15] However, as Autor et al. (1998) note, substantial uncertainty exists concerning the magnitude of the elasticity of substitution. For example, the results of Katz and Murphy (1992) suggest that a two standard error interval around the elasticity of substitution is $[1.01, 2.44]$.

Our estimate of ρ is low when compared to macroeconomic studies that typically assume first-order serial correlation of aggregate productivity between .95 and 1. Recall that $\hat{z}_{it}$ is defined to be the log-deviation of an MSA's exogenous level of productivity from its average. This implies our estimate of ρ is informative only about the persistence of the MSA-specific component of exogenous productivity. Our finding of $\rho = 0.57$ implies that MSA-specific shocks to productivity are much less persistent than shocks to aggregate productivity.

6.3 Effects of Key Modeling Assumptions

We now investigate how two key model ingredients affect our estimates of the model's parameters. We consider how our estimates are affected by assuming intermediate goods are perfect substitutes in the production of the final goods, and by assuming firms do not economize on land when its price is high, that is all land is used in production and counts against the density externality. These are two assumptions imposed by Ciccone and Hall (1996) in their empirical work that we relax in our study.

[15] See also Katz and Murphy (1992), Heckman, Lochner, and Taber (1998), and Krusell, Ohanian, Rull, and Violante (2000).

In all cases, we estimate the parameters by GMM using the procedure of Arellano and Bover (1995) with $\alpha = 0.30$ and $\omega = 0.85$.

Table 4 displays results corresponding to our baseline specification (reproduced from Table 3) and two variations on this specification. Recall that our baseline specification is based on equation (30). To derive (30) we assumed that land is a variable input in production and intermediate goods are imperfect substitutes in production, $\eta < 1$. Suppose we had assumed that intermediate goods are perfect substitutes, $\eta = 1$, but that land is still a variable input into production. In this case the estimation equation is

$$\hat{w}_{eit} = \frac{1}{1-\omega}\frac{\delta-1}{\delta(1-\alpha)}\hat{p}_{it} + \frac{1-\xi}{\xi}\hat{\chi}_{it} + (\xi-1)\,\hat{m}_{it-1} + \hat{z}_{it}. \tag{34}$$

Notice this is identical to the baseline specification except that the output price is dropped. The parameter estimates and standard errors for this case are in the second row of Table 4. Our second variation maintains the assumption $\eta = 1$ and adds to it the assumption that land is a fixed input in the production of intermediate goods. In this case the estimation equation is

$$\hat{w}_{eit} \;=\; \frac{1}{\xi}\left[1-\xi+\frac{\delta-1}{1-\delta\alpha}\right]\hat{\chi}_{it} + (\xi-1)\,\hat{m}_{it} + \left(\frac{\delta-1}{1-\delta\alpha}\right)\hat{u}_{it} + \breve{z}_{it}\ , \tag{35}$$

where $\breve{z}_{it}$ is proportional to $\hat{z}_{it}$. In this case variation in skilled wages is a function of labor market variables only and in particular neither housing rents nor output prices appear. Parameter estimates and asymptotic standard errors for this case are shown in the third row of Table 4.

The second row of Table 4 shows that omitting prices has essentially no impact on our estimate of δ, but raises ξ somewhat, and cuts ρ by more than half. The third row shows that assuming land is fixed in addition to omitting prices leads to much lower estimates of all three parameters. The estimate of δ in this latter case is particularly striking, since it indicates that the density of economic activity is irrelevant for productivity. Clearly our assumption that land is a variable input in production is important to our finding that density has a significant impact on productivity.

7 Implications for Growth

In this section, we quantify the impact of agglomeration on per capita consumption growth. For convenience, we reproduce the balanced growth formula, equation (18), which describes how per capita consumption growth g_c is related to growth in exogenous productivity γ and the growth in the relative price of new land development τ:

$$g_c = \gamma \tau^{\frac{\delta-1}{(1-\alpha)\delta}}.$$

Our quantification strategy is as follows. Given estimates of $\{g_c, \tau, \delta, \alpha\}$ we determine the value of γ implied by the balanced growth equation. We then use this value for γ to determine what the balanced growth equation says consumption growth would have been if agglomeration had no affect on productivity, $\lambda = 1$, so that $\delta = \phi$. In this case consumption growth is given by equation (19). Let g_c^* denote this counterfactual value of g_c and $\hat{g}_c$ denote the empirical estimate of per capita consumption growth. We measure the impact of agglomeration on per capita consumption growth as the percent increase in per capita consumption growth due to agglomeration effects, $\Delta = 100 \times (\hat{g}_c - g_c^*)/(g_c^* - 1)$.

We already have all the ingredients to do this calculation except for an estimate of g_c. We measure consumption as the chain-weighted aggregate of consumption of non-durables and services (exclusive of housing services) plus government consumption plus an estimate of the service flow from the stock of consumer durables.[16] Since our model focuses on the allocation of workers across cities, we use aggregate employment to put consumption in per capita terms. Our measure of aggregate employment is non-farm payroll employment from the BLS. Over the sample period of our panel data, 1985–2006, we estimate $\hat{g}_c = 1.011$.

Table 5 contains our estimates of the impact of agglomeration on per capita consumption growth, Δ. For each indicated parameter configuration we calculate Δ using the calibrated values of $\{\alpha, \tau, \phi\}$ discussed in section 6.1, and the estimates of δ from Table 3 and the surrounding text. The table also includes an estimate of Δ for the case $\delta = 1.055$. This value is in the middle of the range of estimates from the

[16]We obtain measures of the price and quantity of the durables service flow from the Federal Reserve Board. The other components of consumption are from the NIPA.

empirical literature on agglomeration and is close to estimates reported by Ciccone and Hall (1996). In this case we assume $\omega = 0.85$ and use the corresponding values of $\{\alpha, \tau, \phi\}$.

Using our baseline value of $\omega = 0.85$ we find that Δ is 11.7 percent. Assuming $\omega = 0.89$ we find that Δ is 9.6 percent. So, regardless of the assumed value of ω, we find that agglomeration accounts for about 10 percent of per capita consumption growth. Recall that these estimates are based on including the service flow from the stock of consumer durables in our measure of consumption and its price. Our estimates of Δ rise slightly, to 12.1 and 9.7 percent, corresponding to ω equal to 0.85 and 0.89, when we exclude durables.[17] The differences with Table 5 arise from prices for consumer durables rising more slowly than for other consumer goods and services.

It is interesting to compare these estimates to those obtained using a typical value of δ from the literature (the last column of the table). In this case, $\delta = 1.055$, we find that the estimated Δ jumps substantially, to 28 percent. When we exclude the durables service flow from our measures of consumption and its price, this estimate falls slightly, to 27 percent. These findings highlight the sensitivity of our calculations to the exact value of δ and magnify the importance of pinning down the value of this parameter.

Finally, we evaluate the impact of land accumulation in our calculations. Suppose we assume that land in each location is fixed. In this case the derivation of Δ is based on equation (20). When we redo our calculations assuming fixed land we find very similar results: 12.6 and 10.4 percent, corresponding to ω equal to 0.85 and 0.89. It should not be surprising that these numbers change little. Our calibrated τ is close to its upper bound, and at the upper bound the stock of developed land is constant along a balanced growth path.

8 Conclusion

We study the effect of local agglomeration on aggregate growth, providing three main contributions. First, we extend the neo-classical growth model to include location-specific production, consumption, and housing; local developed land as a produced,

[17] This changes the calibration: $\{\tau, \alpha, \phi, \delta\} = \{1.024, 0.20, 0.99, 1.022\}$.

durable input into production and housing; and local agglomeration effects as proposed by Ciccone and Hall (1996). We derive the balanced growth path of this model and show how per capita consumption growth depends on exogenous productivity growth, growth in land prices, and a parameter measuring the net impact on local productivity of agglomeration. Second, we use our model and panel data from U.S. cities to estimate the key agglomeration parameter and find that doubling the density of economic activity in a city raises productivity in that city by 2 percent. This estimate falls below the range of estimates of this parameter in the literature. Third, we quantify the impact of agglomeration on per capita consumption growth by using the balanced growth implications of our model, our estimates of the key agglomeration parameter, and our estimate of land price growth. We find that agglomeration raises per capita consumption growth by about 10 percent.

Many areas of future research are motivated by this paper and we conclude by summarizing a few. First, our finding of a significant contribution of agglomeration to per capita consumption growth needs further consideration both empirically and theoretically. Second, given the important role creation and destruction of cities plays in Rossi-Hansberg and Wright (2007), the robustness of our findings to making endogenous the number of cities should be examined. Third, our methodology for assessing the contribution of agglomeration to growth can be applied to other models, including the Rossi-Hansberg and Wright (2007) model, to quantify the contribution of local human capital externalities to growth. Fourth, our model has many predictions we have not considered but are worth considering. For example, it has predictions for observable features of the cross-section of cities and patterns of migration.

Finally, our model could be useful for understanding how the economy responds to aggregate shocks. A key feature of the business cycle is that residential investment leads non-durable and services consumption and non-residential investment. Fisher (2007) finds that production complementarities between housing and business capital reconcile this observation with business cycle theory. Our model has this essential feature since workers and production establishments are co-located.

References

Acemoglu, D. and V. Guerrieri (2008). Capital deepening and nonbalanced economic growth. *Journal of Political Economy 116*(3), 467–498.

Alvarez, F. and R. Shimer (2008). Search and rest unemployment. Unpublished University of Chicago manuscript.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies 58*(2), 277–297.

Arellano, M. and O. Bover (1995). Another look at the instrumental variables estimation of error-components models. *Journal of Econometrics 68*(1), 29–51.

Autor, D. H., L. F. Katz, and A. B. Krueger (1998). Computing inequality: Have computers changed the labor market? *The Quarterly Journal of Economics 113*(4), 1169–1213.

Basu, S., J. Fernald, J. Fisher, and M. Kimball (2009). Sector-specific technical change. Unpublished Manuscript.

Blundell, R. and S. Bond (2000). Gmm estimation with persistent panel data: An application to production functions. *Econometric Reviews 19*(3), 321–340.

Carlino, G. A. and R. Voith (1992). Accounting for differences in aggregate state productivity. *Regional Science and Urban Economics 22*(4), 597–617.

Ciccone, A. (2002). Agglomeration effects in Europe. *European Economic Review 46*(2), 213–227.

Ciccone, A. and R. E. Hall (1996). Productivity and the density of economic activity. *American Economic Review 86*(1), 54–70.

Ciccone, A. and G. Peri (2006). Identifying human-capital externalities: Theory with applications. *Review of Economic Studies 73*(2), 381–412.

Combes, P.-P., G. Duranton, and L. Gobillon (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics 63*(2), 723–742.

Cooley, T. F. and E. C. Prescott (1995). Economic growth and business cycles. In T. F. Cooley (Ed.), *Frontiers of Business Cycle Research*, pp. 1–38. Princeton University Press.

Davis, M. and J. Heathcote (2005). Housing and the business cycle. *International Economic Review 46*(3), 751–784.

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

Davis, M. and F. Ortalo-Magné (2008). Household expenditures, wages, rents. Manuscript.

Fisher, J. D. (2007). Why does household investment lead business investment over the business cycle? *Journal of Political Economy 115*(1), 141–168.

Greenwood, J., Z. Hercowitz, and P. Krusell (1997). Long run implications of investment-specific technological change. *American Economic Review 78*(3), 342–362.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*(4), 1029–1054.

Heckman, J. J., L. Lochner, and C. Taber (1998). Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics 1*(1), 1–58.

Henderson, J. V. (2003). Marshall's scale economies. *Journal of Urban Economics 53*(1), 1–28.

Holtz-Eakin, D., W. Newey, and H. S. Rosen (1988). Estimating vector autoregressions with panel data. *Econometrica 56*(6), 1371–1395.

Katz, L. F. and K. M. Murphy (1992). Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics 107*(1), 35–78.

Kehoe, T. J., D. K. Levine, and P. M. Romer (1992). On characterizing equilibria of economies with externalities and taxes as solutions to optimization problems. *Economic Theory 2*(1), 43–68.

Kiyotaki, N., A. Michaelides, and K. Nikolov (2008). Winners and losers in housing markets. London School of Economics manuscript.

Krusell, P., L. Ohanian, J. V. R. Rull, and G. Violante (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica 68*(5), 1029–1054.

Lucas, R. E. (2001). Externalities and cities. *Review of Economic Dynamics 4*(2), 245–274.

Lucas, R. E. and E. C. Prescott (1971). Investment under uncertainty. *Econometrica 39*(5), 659–681.

Lucas, R. E. and E. C. Prescott (1974). Equilibrium search and unemployment. *Journal of Economic Theory 106*(3), 514–550.

Ngai, R. and C. Pissarides (2007). Structural change in a multi-sector model of growth. *American Economic Review 97*(1), 429–443.

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

Rosenthal, S. S. and W. C. Strange (2003). Geography, industrial organization, and agglomeration. *The Review of Economics and Statistics 85*(2), 377–393.

Rosenthal, S. S. and W. C. Strange (2004). Evidence on the nature and sources of agglomeration economies. In V. Henderson and J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Volume 4, pp. 2119–2171. North Holland.

Rossi-Hansberg, E. and M. L. J. Wright (2007). Urban structure and growth. *Review of Economic Studies 74*(2), 597–624.

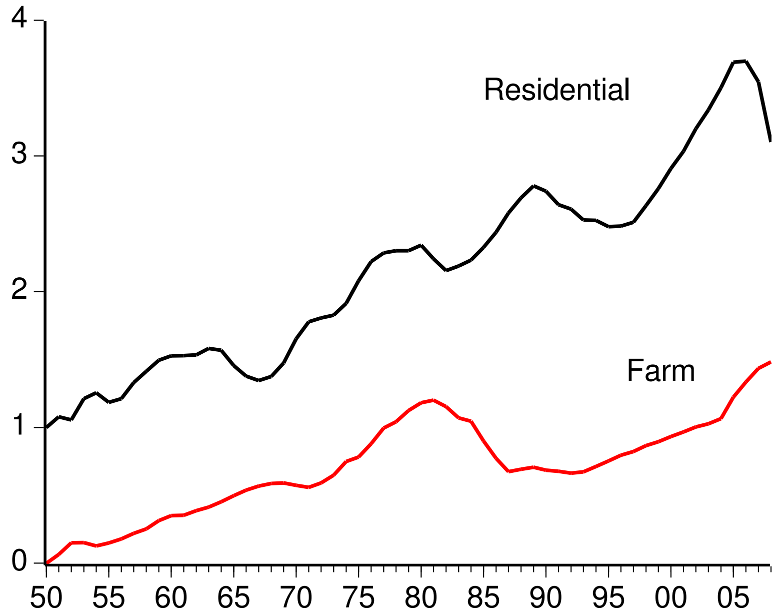
Rozenfeld, H., D. Rybski, X. Gabaix, and Makse (2009). The area and population of cities: New insights from a different perspective on cities. NBER Working Paper No. 15409.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica 26*(3), 393–415.

Van Nieuwerburgh, S. and P.-O. Weill (2009). Why has house price dispersion gone up? Unpublished Manuscript.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step gmm estimators. *Journal of Econometrics 126*(1), 25–51.

Figure 1: Log Real Residential and Farmland Prices, 1950-2008



Source: Davis and Heathcote (2007).

Table 1: Standard Deviations and Correlations in the Data

Variable	Standard Deviation	Correlation with: $\hat{w}_{eit}$	$\hat{p}_{it}$	$\hat{q}_{it}$	$\hat{\chi}_{it}$	$\hat{m}_{it}$
$\hat{w}_{eit}$	0.082	1.00	0.46	-0.03	0.43	0.21
$\hat{p}_{it}$	0.181		1.00	0.49	0.46	0.34
$\hat{q}_{it}$	0.052			1.00	0.23	0.14
$\hat{\chi}_{it}$	0.172				1.00	0.91
$\hat{m}_{it}$	0.224					1.00

Table 2: Unrestricted OLS Regression Coefficients

Dependent Variable: $\hat{w}_{eit}$		Dependent Variable: $\Delta\hat{w}_{eit}$	
$\hat{p}_{it}$	0.185	$\Delta\hat{p}_{it}$	0.148
	(0.039)		(0.103)
$\hat{q}_{it}$	-0.241	$\Delta\hat{q}_{it}$	0.454
	(0.153)		(0.554)
$\hat{\chi}_{it}$	0.571	$\Delta\hat{\chi}_{it}$	0.512
	(0.043)		(0.044)
$\hat{m}_{it}$	-0.349	$\Delta\hat{m}_{it}$	-0.304
	(0.035)		(0.034)
N	484		462
R^2	0.79		0.37

Note: The equation estimated is (30). Standard errors are in parentheses.

Table 3: GMM Parameter Estimates

Parameter	Baseline	$\omega = 0.89$
δ	1.020	1.015
	(0.002)	(0.002)
ξ	0.563	0.563
	(0.004)	(0.004)
ρ	0.569	0.572
	(0.010)	(0.010)
J-test	36.0	36.0
p-value	1	1
m2 test	-0.13	-0.14
p-value	0.45	0.45
m3 test	0.46	0.45
p-value	0.32	0.33

Notes: The equation estimated is (30). Standard errors are in parentheses. The underlying parameter values are discussed in section 6.1. The J-tests have 167 degrees of freedom.

Table 4: Effect of Model Assumptions on GMM Parameter Estimates

Assumption for: Land	Goods	δ	ξ	ρ
Variable	$\eta < 1$	1.020	0.563	0.569
		(0.002)	(0.004)	(0.010)
Variable	$\eta = 1$	1.019	0.607	0.256
		(0.002)	(0.004)	(0.013)
Fixed	$\eta = 1$	0.933	0.524	0.182
		(0.008)	(0.007)	(0.018)

Notes: The equations estimated in the first, second and third rows are (30), (34) and (35). Standard errors are in parentheses.

Table 5: Per Capita Consumption Growth due to Agglomeration

Baseline	$\omega = 0.89$	$\delta = 1.055$
11.7	9.6	28.0

Notes: Table entries are the per cent increase in per capita consumption growth due to agglomeration. The underlying parameter values are discussed in section 6.1.

A Data Appendix

In this appendix, we document how we construct key variables from various data sources and then document how we merge our different data sources together to create our data set for estimation.

A.1 CPS Data on Wages and Hours Worked by Skill

The March 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 1986 through 2007. We choose 1986 as our starting year because the CPS identifies only 15 metropolitan areas in prior years. 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 are 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 in italics)

- Respondent lives in a household, not in group quarters or vacant units ($gq = 1$)
- Is aged 20 to 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 $incwage < 999998$)
- Educational attainment is recorded ($educrec \geq 1$ and $educrec \leq 9$)
- Has an identified metro area of residence (*metarea* non missing)[18]

For each MSA, we create the following three variables:

1. Ratio of skilled labor to unskilled labor, n_{ei}/n_{ui}
2. Ratio of total wages paid to total wages paid to low skill workers, χ_i
3. Average weekly wage of skilled workers, $w_{e,i}$.

[18] According to notes from the IPUMS-CPS, the metro area of residence was not collected from respondents, but added by the Census Bureau. The metro areas of residence is based on FIPS codes used in the 1990 census.

We use the *educrec* categorical variable to label respondents as either "unskilled" or "skilled" workers. Skilled workers are assumed to have completed 4 years of college. Everyone else in the sample is assumed to be an skilled worker.

The variable n_{ei} is created as the total of weeks worked the previous calendar year ($wkswork1$) multiplied by the number of hours per week the respondent usually worked ($uhrswork$) for skilled workers. The variable n_{ui} is created as the same product, but for unskilled workers. For each respondent, we weigh the product of $wkswork1$ and $uhrswork$ using the IPUMS-CPS sampling person weights, $perwt$.

χ_i is computed as

$$\frac{w_{e,i}n_{ei} + w_{u,i}n_{ui}}{w_{u,i}n_{ui}} = \frac{\sum_{j\epsilon MSA_i} perwt_j \cdot wages_j}{\sum_{j\epsilon MSA_i} perwt_j \cdot wages_j \cdot 1\{unskilled_j\}}$$

for respondent j in MSA i, i.e. as the sum of all workers' pre-tax wage and salary income for the previous calendar year ($incwage$) divided by the sum of all unskilled 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.

$w_{e,i}$ is created as the sum of all skilled workers' pre-tax wage and salary income for the previous calendar year (created as an input into χ_i) divided by n_{ei}.

A.2 BEA Data on Output Prices

We use two data sources from within the BEA web site: The Annual Industry Accounts, *http://www.bea.gov/industry/index.htm#annual*, and the Regional Economic Accounts data on Local Area Personal Income, *http://www.bea.gov/regional/reis/*.

Chain-type price indexes for industry output are available over the 1947-2007 period in the Annual Industry Accounts. Many industry price indexes are missing in 2007, so we do not use data from that year. 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. Earnings is inclusive of wage and salary disbursements, supplements to wages and salaries, and proprietors' income.

Thus, 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.

We now describe our measurement in detail. For this we use notation that does not correspond with that in the main text. Denote $g_{t,j}$ as the growth rate of the price of industry output j from periods t to $t+1$ and 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

$j = 1, \ldots, N$ 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

$$g_t^i \;=\; \sum_{j=1}^{N} \omega_{t,j}^i \; g_{t,j} \; . \tag{36}$$

The weight on each industry, $\omega_{t,j}^i$, is the share of total MSA earnings attributable to earnings of industry j in MSA i in year t:

$$\omega_{t,j}^i \;=\; \frac{\epsilon_{t,j}^i}{\sum\limits_{k=1}^{N} \epsilon_{t,k}^i} \; , \tag{37}$$

where $\epsilon_{t,j}^i$ stands for total earnings of employees in industry j in MSA i during year t. In these computations, we only consider earnings from non-farm private industries. 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_t^i.

Ideally, we would compute the growth rate of the price of output produced in city i between years t and $t+1$ as

$$\sum_{j=1}^{N} \phi_{t,j}^i \; g_{t,j} \; , \tag{38}$$

with $\phi_{t,j}^i$ equal to the fraction of the nominal value of output in year t in MSA i that is accounted for by industry j. In an environment in which (a) output in each industry is produced by a set of identical firms all using a Cobb-Douglas combination of capital, labor, and land and (b) the labor-share of output is identical in each industry, assumptions that hold in our model, then industry j's share of nominal GDP in MSA i in year t, $\phi_{t,j}^i$, is equal to its earnings share $\omega_{t,j}^i$, and equations (36) and (38) are equivalent. In these calculations, we assume that proprietors' income are payments to labor.[19]

A few details complicate these calculations. 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 zero.[20] Also, some of the industry-MSA-year employment estimates are marked

[19] In the event that proprietors' income includes some payments to capital, equations (36) and (38) are equivalent as long as capital's share of proprietors' income and the fraction of earnings attributable to proprietors' income are both constant across industries.

[20] The three reasons that are listed for omission are (a) avoid disclosure of confidential information (code D), (b) earnings are less than $50,000 (code L), or (c) data not available for this year (code N). These omissions occur in approximately six percent of industry-MSA-year cells from 1969 to the mid-1990s and about thirteen percent of cells from the mid-1990s through 2006.

with code E. According to the BEA web site, 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 spanning the years 2001-2006.

We map SIC and NAICS industry employment from Tables CA05 and CA05N to prices from the Annual Industry Accounts according to the tables below. The tables below list all the categories of nonfarm private employment. The sum of the earnings estimates in each of these categories is considered as total nonfarm private earnings, and is used to compute the denominator of equation (37).

Data for Earnings Weights, $w^i_{t,j}$ Regional Accounts Table CA05, 1969-2000		Data for Growth in Prices, $g^p_{t,j}$ Industry Accounts, 1969-2001	
Line	Label	Line	Label
100	Agricultural services, forestry fishing and other	3	Agriculture, forestry, fishing and hunting
200	Mining	6	Mining
300	Construction	11	Construction
400	Manufacturing	12	Manufacturing
500*	Transportation and public utilities less electric, gas, and sanitary services	36	Transportation and warehousing
570	Electric, gas, and sanitary services	10	Utilities
610	Wholesale trade	34	Wholesale trade
620	Retail trade	35	Retail trade
700	Finance, insurance and real estate	50	Finance, insurance, real estate, rental and leasing
800	Services	59	Professional and business services

Data for Earnings Weights, $w^i_{t,j}$ Regional Accounts Table CA05N, 2001-2005		Data for Growth in Prices, $g^p_{t,j}$ Industry Accounts, 2001-2006	
Line	Label	Line	Label
100	Forestry, fishing, related activities and other	5	Forestry, fishing and related activities
200	Mining	6	Mining
300	Utilities	10	Utilities
400	Construction	11	Construction
500	Manufacturing	12	Manufacturing
600	Wholesale trade	34	Wholesale trade
700	Retail trade	35	Retail trade
800	Transportation and warehousing	36	Transportation and warehousing
900	Information	45	Information
1000	Finance and insurance	51	Finance and insurance
1100	Real estate and rental and leasing	56	Real estate and rental and leasing
1200	Professional, scientific and technical services	60	Professional, scientific and technical services
1300	Management of companies and enterprises	64	Management of companies and enterprises
1400	Administrative and waste services	65	Administrative and waste management services
1500	Educational services	69	Educational services
1600	Health care and social assistance	70	Health care and social assistance
1700	Arts, entertainment and recreation	75	Arts, entertainment and recreation
1800	Accommodation and food services	78	Accommodation and food services
1900	Other services except public administration	81	Other services except government

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 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).

A.3 BLS and 1990 Decennial Census of Housing Data on Housing Rents

We create annual estimates over the 1985-2006 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 because the metropolitan area of residence is identified for many more metropolitan areas than in the 2000 DCH.

With IPUMS-USA variables in italics, we restrict the 1990 DCH sample to renter non-farm 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 (> 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 employing hedonic regressions.

In the second step, we extrapolate the annual rental price of housing in each metro area forward from 1990 to 2006 and backwards from 1990 to 1985 using annual MSA-specific constant-quality price indexes for rental units. These price indexes for tenant rents 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 extend back to 1985 and we exclude these from our sample.

After merging the 1990 DCH with the BLS price indexes, and eliminating the MSAs for which price indexes are not available back to 1985, we are left with annual estimates of the rental price of housing units over the 1985-2006 period for 24 MSAs. These MSAs are: Anchorage, AK; Atlanta, GA; Boston, MA; Chicago, IL; Cincinnati, OH; Cleveland, OH; Dallas, TX; Denver, CO; Detroit, MI; Honolulu, HI; Houston, TX; Kansas City, MO; Los Angeles, CA; Miami, FL; Milwaukee, WI; Minneapolis, MN; New York, NY; Philadelphia, PA; Pittsburgh, PA; Portland, OR; San Diego, CA; San Francisco, CA; Seattle, WA; St. Louis, MO.

A.4 Merging the Data

To create our data set, we merge the CPS data on wages and employment (section A.1) with the BEA data on output prices (A.2) and the annual data we construct on housing rents by merging the BLS rental price indexes with information on housing rents in the 1990 Decennial Census of Housing (section A.3). After all data are merged, we are left with a balanced panel of 22 MSAs over 22 years, 1985-2006. Anchorage, AK and Portland, OR are excluded from our data due to lack of wage and employment data from the CPS. Over our sample period, 36 percent of the population of the United States live in one of the the 22 MSAs in our study.

In every MSA and date, the minimum number of respondents from the CPS is never less than 200. The median number of respondents is about 540 until about 2000, at which point the median jumps to about 1,000. The maximum number of respondents is always above 3,500 and is typically about 5,000.

All data files are merged by MSA. Note that the MSA definitions are not completely consistent across data sets. The MSA definitions in the CPS data are consistent with the definitions as of the 1990 Census. In the BEA data, MSAs definitions are given by the list in the November, 2007 report of the Office of Management and Budget (OMB).[21] The MSA definitions in the BLS rent indexes are also based on the definitions in OMB, but the definitions can change over time (as the OMB changes its definition of MSAs over time); further, the rent indexes for New York, Los Angeles, and Chicago are based on the concept of the Consolidated Metropolitan Statistical Area (CMSA). CMSAs include more counties than simple MSAs.

B Monte Carlo Study

We perform a Monte Carlo study using simulated data whose distribution closely approximates that of our own data. First, we generate a normal $N(0,1)$ value for z for the first time period. We then set each of the right-hand side variables (p^h, q, χ, and m) equal to a normal random variable plus $0.5 * z$. We then update z using the specified AR(1) process (using the appropriate value for ρ) and repeat.

Once we generate the entire panel, we restandardize the right-hand side variables so that the have the same first and second moments as our data, and then generate $w_{ei,t}$ using equation (30). In this process, we rescale the variance of z such that the simulated variance of $w_{ei,t}$ is the same as in the data. This procedure allows z to be correlated with the right-hand side variables, but also sets lagged values of the right-hand side variables orthogonal to the innovations to z, which is the maintained assumption of the estimator.

We repeat this procedure 10,000 times (thus generating 10,000 data sets) for each of the following four combinations of parameter values: $\delta = \{1.011, 1.040\}$,

[21] For a complete list of the counties comprising each MSA, go to *http://www.census.gov/population/www/metroareas/metrodef.html*.

$\rho = \{0.60, 0.95\}$. In all cases, we set $\xi = 0.60$, $\omega = 0.85$ and $\alpha = 0.30$. We report the results for an unrestricted weight matrix, in which the residuals from different time periods are down-weighted with the Bartlett kernel. We use this method when we report parameter estimates and standard errors using the actual data in the main body of the paper.[22]

In table 6, for each combination of parameters we report six statistics. Conceptually, the first three statistics we report—the average estimate of the parameter, the mean absolute deviation of the parameter estimate, and the root mean square error of the parameter estimate—are indicative of potential biases to the parameter estimate. The other three statistics we report are indicative of potential bias to estimated standard errors: the probability of rejecting the null (which should equal 0.05 if the test is correctly sized), the 95 percent coverage probability, and a scale factor (to be defined). The 95 percent coverage probability is the percentage of the simulations in which the distance between the parameter estimate and the true parameter is in the 95 percent confidence interval. The scale factor represents the adjustment to reported standard errors such that the probability of rejecting the null is 0.05. That is, if the reported standard error is S and the scale factor is F, the probability of rejecting the null when the standard error is of size $S * F$ is 0.05.

Based on the results in this table, we draw two main conclusions. First, our estimates of δ, ρ, and ξ are, for all practical purposes, unbiased. Second, the asymptotic standard errors are too small for all parameter estimates: The probability of rejecting the null is always larger than 0.05 and the 95 percent coverage probability is always less than 0.95. For the case of $\rho = 0.60$, the scale factors suggest that our standard errors are too small by a factor of 2.35 for δ, 1.44 for ρ, and 2.09 for ξ.

[22] We have also performed the Monte Carlo for using a "restricted" weight matrix in which cross sectional residuals from different time periods are not allowed to be correlated. This is the original formulation of Holtz-Eakin, Newey, and Rosen (1988) that results in a block-diagonal weight matrix. The results are very similar and are available on request.

Table 6: Monte Carlo Results

	$\rho = 0.60$			$\rho = 0.95$		
	δ	ρ	ξ	δ	ρ	ξ
			$\delta = 1.011$			
Expected Value	1.0117	0.6030	0.5975	1.0112	0.9406	0.5994
Mean Absolute Deviation	0.0008	0.0119	0.0032	0.0002	0.0117	0.0009
Root Mean Square Error	0.0009	0.0152	0.0039	0.0002	0.0148	0.0011
Prob. Reject Null	0.49	0.17	0.37	0.37	0.28	0.28
95% Coverage Prob.	0.51	0.83	0.63	0.63	0.72	0.72
Scale Factor	2.35	1.44	2.09	2.02	1.75	1.83
			$\delta = 1.040$			
Expected Value	1.0407	0.6030	0.5975	1.0402	0.9406	0.5994
Mean Absolute Deviation	0.0008	0.0119	0.0032	0.0002	0.0117	0.0009
Root Mean Square Error	0.0010	0.0152	0.0039	0.0003	0.0148	0.0011
Prob. Reject Null	0.49	0.17	0.37	0.37	0.28	0.28
95% Coverage Prob.	0.51	0.83	0.63	0.63	0.72	0.72
Scale Factor	2.35	1.44	2.09	2.02	1.75	1.83