Agglomeration and Sorting

Yujiang Chen and Coen Teulings*

January 2020

[Preliminary]

Abstract

Recent papers suggest a strong interaction between agglomeration externalities and human capital. We develop a general equilibrium model with multiple regions where agglomeration benefits are increasing in human capital. Regions can either be organized as cities with a CBD or as rural areas. The city form is conducive to knowledge spill-overs but city size is limited by commuting cost. We estimate the model on US data on housing prices and wages for 47 states and 34 metropolitan areas from 1979 till 2015. We find strong support for the model. We use the model for the calculation of two counterfactuals: first without cities, and second without any agglomeration benefits. We find that land would loose half its value without cities and almost all its value without any agglomeration benefits.

JEL classification: J24, J31, I26, R12, R13

Keywords: Spatial Sorting, Agglomeration externalities, Cities, Regional wage differentials, Regional house prices

*Yujiang Chen, University of Cambridge, yc349@cam.ac.uk; Coen Teulings, Utrecht University and University of Cambridge, c.n.teulings@outlook.com. We thank Daron Acemoglu, Gabriel Ahlfeldt, David Albouy, Michael Amior, David Autor, Nicholas Bloom, Jan Eeckhout, Pieter Gautier, Edward Glaeser, Esteban Rossi-Hansberg, William Janeway, Koen Jochmans, Alan Manning and Guy Michaels, for helpful comments. We also benefited from very useful feedback from seminar audiences at Cambridge, CREST, CSIC, IFS, MIT, NBER, Utrecht and UEA. All mistakes remain our own.

1 Introduction

Around 1920, Frits Philips was pondering where to set up the new factory of electric light bulbs. He considered several villages in the South Eastern part of the Netherlands, like Helmond, Veghel, and Veldhoven. He ended up in building his factory in the village of Eindhoven. Subsequently, this little village went through several decades of exceptional growth. By 1950, Eindhoven was the 7th city of the Netherlands, and 20 years later it had climbed to the 5th rank, a position Eindhoven still holds. Philips Electronics built extensive laboratory facilities, which were renowned in the industry. The city started its own technical university. From 1970 onwards, Philips Electronics went through a difficult episode. It had a hard time marketing its excellent technological innovations and went almost bankrupt. The renowned laboratories were closed down. Eventually, Philips decided to move its headquarters to Amsterdam, seeking a more open labour market and a better connection to the outside world. Eindhoven experienced a deep trough. But in the end, the backbone of former researchers of Phillips' laboratories, well trained engineers, many of them receiving their education at Eindhoven's technical university, saved the city. There were many new startups, often supported by Philips. Nowadays, the city is striving again, hosting ASML, the world leader in new production technologies for ICs. The current market capitalization of ASML exceeds that of its parent Philips.

This story is just one of many. Gennaioli, La Porta, Lopez-de-Silanes and Shleifer (2013) analyze the regional distribution of human capital and economic activity in more than 100 countries across the world. They show that in each country human capital tends to cluster in a small number of regions. GDP per capita in these regions is much higher than the nation-wide mean. A simple regression of the regional GDP per capita on the regional mean years of education yields returns to an additional year of education of e.g. 54% for Brazil, 31% for India, 23% for Colombia, and 55% for Russia, higher than any reasonable estimate of the private return to human capital, although one might have reservations regarding the causality of these returns. However, their results suggest that there are either large agglomeration externalities of human capital or that some regions have inherent features that makes them more productive than others, see Moretti (2004) for an overview.

We develop a general equilibrium model with multiple regions where agglomeration benefits are increasing in human capital. Regions specialize in a particular activity, that is characterized by the mean level of human capital that this activity requires. Based on the strong persistence in the industrial structure of regions, see e.g. Amior and Manning (2018), we treat the human capital requirement of a region as largely exogenously fixed in the short and intermediate run. Following Gennaioli et.al. (2013), we assume that the benefits from the exchange of ideas are increasing in the mean level of human capital. This exchange of ideas cannot be adequately internalized by employers. Hence, wages for workers with equal human capital are higher in regions with a high average level of human capital. Since there is free interregional labour mobility, these wage differentials are offset by the interregional variation in the cost of housing services, or more precisely, in the price of land on which these houses are build.

To analyze the extent of spatial misallocation, we add a distinction between two forms of spatial organization for a region: cities versus rural areas. The city form is modeled as in Lucas and Rossi-Hansberg (2002) and Rossi-Hansberg and Wright (2007). Cities are organized around a CBD that for simplicity is assumed not to use any land, as in Rossi-Hansberg and Wright 2007), and that is surrounded by residential areas. While workers have to commute to the CBD, ideas don't have to travel but can reside in the CBD where they can be exchanged without spatial transmission cost between all workers who work at that location. This structure is by construction finite since beyond a certain distance from the CBD, commuting becomes too costly. To the contrary, the structure of rural areas is scale free. Workers live and work at the same location. Hence, they do not have to commute. Instead, ideas have to travel between the job locations of workers. This leads to lower commuting cost, but less exchange of ideas since the spatial transmission of ideas is costly. Since this rural form has no spatial differentiation between a CBD and its surrounding suburbs, all location are identical and hence the rural form can be infinitely extended. Since high human capital activities benefit more strongly from the exchange of ideas, they have a comparative advantage in cities where the CBD facilitates a free exchange of ideas.

This model yields a rich set of predictions. The mechanisms at work can be nicely illustrated by effect of an exogenous amenity like January temperature on house prices. In a general equilibrium model without agglomeration benefits, a high January temperature would affect only regional land prices, not wages. In order to offset the pull-force of the agreeable climate on labour supply, land prices would be higher. In our model with agglomeration benefits, there is also an effect on wages: high land prices reduce per capita land use as workers substitute away to other consumption. A lower land use increases the population density in region and therefore the agglomeration benefits. This leads to higher wages which in turn drive up house prices even further. The model predicts agglomeration benefits to be more sensitive to the mean level of human capital in cities than in rural areas because cities have an additional margin of adjustment: cities can adjust average land use per capita and the city-size (= radius), while rural areas can only adjust land use.

The model is estimated on US data on wages and house prices for 47 states and 34 metropolitan areas from 1979 till 2016. The estimation of the regional fixed effect in log wages and the average level of human capital in the region requires special attention due to unobserved human capital. We apply a Proportionality Assumption derived from a Single Index Model for human capital, see e.g. Teulings (1995). We provide empirical evidence in favour of this assumption, using occupation dummies as an instrument for human capital. We argue that the Proportionality Assumption provides an upper bound for the correlation between observed and unobserved human capital.

By the same argument, it provides a lower bound for the regional fixed effect in log wages. Using this correction, we find that for each 10% increase in regional wages due to the private return to human capital, there is an additional 5% increase in wages due to knowledge spill-overs.

Empirically, the human capital requirement of a region is indeed an important determinant of this agglomeration effect, consistent with the model's predictions. We treat this human capital requirement as exogenous. However, In our empirical application, we must allow for the partial endogeneity of the mean level of human capital. Hence, we use Bartik instruments based on the nation-wide shifts in the industry composition and in the mean level of education in each industry. These instruments turn out to be very strong. The model yields multiple over-identifying restrictions, in particular due to the joint predictions for the effects on wages and house prices. By and large, we find strong support for the model. Our estimate of the agglomeration elasticity of human capital is similar to in Genniaoli et.al. (2013). Since the city-form is more conducive to knowledge spillovers and since these spillovers are more important for activities with a high human capital requirement, cities have a comparative advantage in these activities. Though we treat this human capital requirement as exogenous in our model, it will be endogenous in the very long run, compare Desmet and Rossi-Hansberg (2009). Hence, high human capital activities are expected to be located in cities. This prediction is strongly supported by the data.

Recently, Herkenhoff, Ohanian, and Prescott (2018) and Hsieh and Moretti (2019) have argued that incumbents have limited the access to high productive areas by means of land use restriction in order to preserve the value of their housing property. This is supposed to have led to substantial spatial misallocation of labour. In this paper, we argue that this claim ignores the limits in extending the size of high productive areas (usually cities) imposed by the cost of commuting. We find the several rural areas with relatively high human capital requirements would benefit from the transition to the urban form of spatial organization. Implementing this transition would require some form of collective action.

This raises a more fundamental issue. Can incumbents in a region benefit from limiting the supply of housing to protect the scarcity value of their property, e.g. by zoning restrictions? A standard supply and demand logic suggests so. However, in a model with agglomeration benefits, this is less evident. A larger population increases the size agglomeration benefits and will therefore push up wages and hence land prices: the housing demand curve is upward sloped. Incumbents therefore have no incentive to restrain the construction of new housing. If anything, incumbents have an incentive to engage in collective action to subsidize new construction paid for by a Henry George taxation on the excess landvalue to benefit from the positive externality, as in Rossi-Hansberg and Wright (2007). However, we assume that land owners are unable to solve this collective action problem. Instead, we assume that the market for residential land is fully decentralized, both in cities (as in Lucas and Rossi-Hansberg 2002) and in rural areas.

We use the model for the calculation of a number of counterfactuals, first, transforming current cities into rural areas, and second, setting all agglomeration benefits on the production side equal to zero. We calculate these counterfactuals under two assumptions, either with an infinitely elastic nation wide labour supply, hence keeping workers' utility constant, or with a fixed labour supply and hence with an endogenous worker utility. We find that land looses half of its value when we rule out the city form and almost all of its value if there are no agglomeration benefits at all.

The structure of the paper is as follows. Section 2 discusses our theoretical model. Our empirical evidence is presented in Section 3. In Section 4, we use our model for the calculation of counterfactuals. Section 5 concludes.

2 Spatial Equilibrium Model

2.1 General structure

We consider an economy consisting of regions indexed r. Workers are endowed with a level of human capital h. Each worker supplies one unit of labour. Her wage is her only source of income. For simplicity, there is no physical capital in this economy. Production in each region is governed by a Leontief technology requiring the inputs of all levels of human capital h in fixed proportions. These proportions are given by the density function of a normal distribution with variance σ^2 common to all regions. However, the mean of this distribution H_r varies between regions, depending on the type of tradable commodity that is produced in the region; H_r is exogenous. Tradeables are traded on the nation-wide commodity market. Since all workers have the same homothetic utility function and since there are no transport cost of tradeables between regions, the composition of consumption is the same across regions. All land rents are earned by a class of absentee landlords. Each region has three exogenous characteristics: the mean level of human capital H_r, an exogenous consumption amenity (January temperature T_r), and whether the region is organized as a rural area or a city. There is perfect competition on national and regional labour and product markets, and on regional land markets (a national land market does not exist, since land is non-tradable between regions).

The model consist of three building blocks:

1. workers' utility function: free interregional labour mobility sets the utility of a worker with human capital h equal to some exogenous benchmark for that level of human capital;

2. regional housing markets: workers choose the lot size of their house as to maximize their utility. Regional log land prices p_r adjust to clear the land market. Competition between regions drives land prices up or down to the point where workers are indifferent between regions;

3. agglomeration externalities: depending on the mean level of human capital and on spatial form (rural areas versus cities), a region benefits from agglomeration externalities.

These three blocks will be discussed in the next subsections. In equilibrium, the log wage $w_r(h)$ of a worker with human capital h living in region r will be a linear function of her human capital. Since we have not defined the units of measurement of h, the slope of this function can be normalized to unity without loss of generality by a proper choice of the measure of human capital.

$$w_r(h) = \omega_r + h. \tag{1}$$

The intercept ω_r differs between regions and is determined endogenously. The equilibrium to this economy can be described by relations for the endogenous region specific variables ω_r and p_r that are (almost) linear in the exogenous region specific variables H_r and T_r. Without loss of generality, the nation-wide means of these aggregate variables are normalized to zero

$$\mathrm{E}[H_r] = \mathrm{E}[T_r] = \mathrm{E}[\omega_r] = \mathrm{E}[p_r] = 0. \tag{2}$$

Since their variances are small, the use of a first order Taylor approximation at the point $H_r = T_r = \omega_r = p_r = 0$ is justified for most relations in the model. Whenever we use the symbol $\cong$, we refer to a Taylor approximation.

2.2 Workers' utility

Workers choose in which region r to live and work at the beginning of their career. They do so as to maximize their utility. Interregional mobility is free. Worker mobility will therefore equalize the utility of each h-type worker across regions. Since human capital enhances her earning capacity, this nation-wide benchmark utility $u(h)$ depends on the human capital of the worker. We assume that total labour supply is perfectly elastic: workers migrate in and out the country until the utility offered to workers in the United States is equal to the utility elsewhere. The assumption implies that we can treat $u(h)$ as fixed function. Hence, workers' utility remains unaffected and all welfare gains or losses fall upon the class of landlords. We make this assumption just for the sake of convenience. It does not affect our empirical inference in Section 3. For the welfare analysis in Section 4 this assumption is relaxed. Without loss of generality, this exogenous benchmark utility is normalized to the human capital index h:

$$u(h) = h. \tag{3}$$

Workers derive utility from the private consumption of tradeables and non-tradeables and from the availability of amenities/public goods. Tradeables are traded across regions at a constant nation-wide log price p, which is treated as the numeraire: $p = 0$. The non-tradable consumption good is land, either directly, land that is used for residential purposes, or indirectly, land that is used e.g. for shopping malls, where the price of the merchandise reflects cost differentials due to variation in the price of land, or the land that is used by workers providing non-tradable services and who get compensated for the higher land prices by higher wages.

We consider two types of regional amenities: the January temperature T_r and the spatial density of income z_r. People prefer to live in regions where January temperature is more agreeable, see e.g. Glaeser (2009). Ahlfeldt, Redding, Sturm and Wolf (2015) show in their study of Berlin that there are strong and highly localized agglomeration externalities in residential areas. A high local density allows a dense network of services like shops, restaurants, and cultural performances to be sustained, as in De Groot, Marlet, Teulings and Vermeulen (2015) and Diamond (2016). Hence, the higher the log income per unit of land, z_r, the higher the quality of the network of these services; z_r will be endogenously determined. The private benefits of regional amenities cannot be priced directly. Hence, they are reflected in regional house prices. As will be shown, these externalities play a crucial role in explaining the data.

Workers' utility function is homothetic with a constant elasticity of substitution between tradeables and land. The equilibrium condition for interregional migration can be derived from the cost function that corresponds to a CES utility function

$$\begin{aligned} \underbrace{w_r(h)}_{\text{income = cost}} &= \underbrace{h}_{\text{benchmark utility}} + \underbrace{\overline{\eta}^{-1} \ln\left(\overline{\lambda} + \lambda \mathrm{e}^{\overline{\eta} p_r}\right)}_{\text{price index}} - \underbrace{\alpha' a_r}_{\text{public goods}} \\ &\cong h + \lambda p_r - \alpha' a_r \end{aligned} \tag{4}$$

where $a_r \equiv [T_r, z_r]'$ is the vector of public goods, and where λ is the land share in expenditure when the price of land is unity, $p_r = 0$, and where η is the elasticity of substitution between tradeables and land; in accordance with the empirical evidence, we assume $0 < \eta \leq 1$ and $0 < \lambda < \frac{1}{2}$. We adopt the convention that a bar on top of a parameter denotes its complement with respect to unity, so $\overline{\lambda} \equiv 1 - \lambda$.[1] The left hand side is log income (= cost of obtaining a utility level h). The first term on the right hand side is the benchmark utility level. This term cancels against h on the left hand side, showing the consistency of our specification of the endogenous function $w_r(h)$ in equation (1) with the exogenous benchmark utility function in equation (3). The second term is the CES price index in region r. Since the price of tradeables is normalized to unity, it drops out. The final term measures the compensating differentials for regional amenities. Other things equal, regions with high amenities will have lower cost for offering a utility level h. The parameter

[1] In the second line, we use

$$\ln\left(\overline{\lambda} + \lambda e^{\overline{\eta} p}\right) = \overline{\eta} \lambda p + O\left(p^2\right).$$

vector $\alpha \equiv [\alpha_T, \alpha_z]'$ measures the compensating differential for one unit of the amenity as a share of disposable income. Equation (4) must hold identically for all h. This yields an approximately linear equation with unit slope for the worker's human capital h as in equation (1), where the intercept ω_r satisfies

$$\omega_r \cong \lambda p_r - \alpha' a_r. \tag{5}$$

Log wages in region r must compensate for the level of log land prices multiplied by the land share in expenditure and for the availability of amenities. Interregional labour mobility therefore imposes that log wages depend negatively on amenities, controlling for the log price of land p_r.

2.3 Regional land markets

By Shephard's Lemma, the partial derivative of the cost function with respect to price of land is equal to the demand for land. This implies that the partial derivative of the log cost function with respect to the log price of land is equal to the land share in expenditure.[2] Taking logs yields:

$$\begin{aligned} \ln\left[\partial e\left(0, p_r, h\right)/\partial p_r\right] &= \ln\lambda + (1-\eta)\, p_r - \ln\left(1 - \lambda + \lambda \mathrm{e}^{(1-\eta)p_r}\right) \\ &= l_r\left(h\right) + p_r - w_r\left(h\right), \end{aligned}$$

where the second line is the definition of the log land share in expenditure. Some simplification and applying a Taylor approximation yields

$$\begin{aligned} l_r\left(h\right) &\cong \ln\lambda - \lambda_\eta p_r + w_r\left(h\right), \\ \lambda_\eta &\equiv 1 - \overline{\lambda}\overline{\eta}; \quad 0 < \lambda_\eta \leq 1. \end{aligned} \tag{6}$$

For either $p_r = 0$ (the nation-wide mean of house prices) or $\eta = 1$ (Cobb-Douglas utility: $\lambda_\eta = 1$), the log spending on land is equal to log total expenditure $w_r\left(h\right)$ plus the log land share in this expenditure, $\ln\lambda$. Substitution of equation (1) and (5) yields an expression for the average log land use l_r in region r as a function of the regional fixed effect in the wage equation ω_r, the regional

[2] Let $E\left(P_c, P, U\right)$ be expenditure as a function of the price of tradables P_c, the price of land P, and the utility level u. Since $P_c = 1$, Shephard's lemma implies

$$\begin{aligned} &\partial E\left(1, P, u\right)/\partial P = L, \\ &\partial e\left(0, p, u\right)/\partial p = \frac{\partial E\left(1, P, u\right)/\partial P}{E\left(1, P, u\right)}\frac{\partial P}{\partial p} = \frac{L \cdot P}{E\left(1, P, u\right)}, \end{aligned}$$

where $e\left(0, p, u\right) \equiv \ln E\left(1, \mathrm{e}^p, u\right)$. The final expression is the land share in total expenditure.

amenities a_r, and the regional mean level of human capital H_r[3]

$$l_r \equiv l_r(H_r) \cong \ln\lambda - \frac{\overline{\lambda}}{\lambda}\eta\omega_r - \frac{\lambda_\eta}{\lambda}\alpha' a_r + H_r. \tag{7}$$

Land use depends negatively on regional wages and amenities and positively on the mean level of human capital. High wages raise the attractiveness of a region, which pushes up the regional price of land. Then, the substitution effect reduces average land use. Similarly, high amenities make a region attractive and therefore increase the price of land, which also reduces average land use. Since workers' utility function is homothetic, average land use is proportional to disposable income and hence to the mean level of human capital.

The density of income in region r is the average income per worker divided by the average land use, in logs

$$\begin{aligned} z_r &= w_r(H_r) - l_r \cong \lambda_\eta\lambda_z(\omega_r + \alpha_T T_r), && (8)\\ \lambda_z &\equiv (\lambda - \lambda_\eta\alpha_z)^{-1}, \end{aligned}$$

where the second equality in the first line follows from substitution of the equations (6) and (5) and where we omit the constant for the sake of convenience. We assume $\lambda_z > 0$, which holds for the parameters we derive from the literature, see Section 3.4. Substitution in equation (7) yields

$$\begin{aligned} l_r &\cong -\lambda_\omega\omega_r - \lambda_\eta\lambda_z\alpha_T T_r + H_r, && (9)\\ p_r &\cong \lambda_z(\omega_r + \alpha_T T_r),\\ \lambda_\omega &\equiv \lambda^{-1}\left(\overline{\lambda}\eta + \lambda_\eta^2\lambda_z\alpha_z\right) > 0, \end{aligned}$$

again omitting constants. Compared to equation (7), the consumption agglomeration effect z_r increases the effect of ω_r and T_r on log land prices.

2.4 Agglomeration externalities

Intra-regional agglomeration externalities or knowledge spillovers are modelled similar to Gennaioli, La Porta, Silanes and Shleifer (2013). Our specification reads

$$\omega_r = \psi(\theta H_r + m_r) + \omega_0, \tag{10}$$

where m_r is the number of workers that contribute to the agglomeration externality at a particular point in space, and where ψ and θ are weakly positive parameters. For $\psi = 0$, there are no

[3]We use the mean log value and the log mean value of a house interchangeably. This is incorrect by Jensen's inequality. Due to the normality of the distribution of h, the difference is a constant, $\frac{1}{2}\sigma^2$. We treat these terms as normalizing constants.

knowledge spillovers. For $\theta = 0$, knowledge spillovers depend only on the number of workers, not on the level of human capital. Due to the normalization of the average return to human capital to unity, $\theta = 1$ would imply that knowledge spillovers are proportional to the total wage bill. Gennaioli et.al. (2013) report evidence that knowledge spillovers increase more than proportional to the average level of human capital of the regional workforce. This is the case if $\theta > 1$. The parameter is ω_0 is a mean shifter; we set it equal to zero in this section, but we shall need it as mean shifter of ω_r when we confront the model to the data in the next section.

Our modeling of the intra-regional spatial structure combines ideas from Lucas and Rossi-Hansberg (2002), Rossi-Hansberg and Wright (2007), and Ahlfeldt, Redding, Sturm and Wolf (2015). In the cities considered in Lucas and Rossi-Hansberg and Ahlfeldt et.al. workers have to commute between their home and job location and ideas have to travel between the locations of different jobs. We consider two opposite archetypical spatial structures: rural areas and cities. In a rural area, people work at their home location. Hence, workers do not commute and ideas have to travel. In a city, it is exactly the opposite: jobs are concentrated in a Central Business District (CBD). Hence, ideas don't travel, but workers have to commute. We discuss both archetypes below.

2.4.1 Rural areas

In rural areas, workers work at the same location as they live and all h-type workers are spread homogeneously across space. Like Lucas and Rossi Hansberg (2002) and Ahlfeldt et.al. (2015), the travel of knowledge spills across space comes at a cost: at distance s, only a fraction $1 - \delta s$ of the spillover survives. The maximum distance ideas can travel is therefore δ^{-1}. Only workers working within a distance δ^{-1} contribute to the knowledge spillover for a particular worker. Hence, the knowledge spillover ω_r^r in region r (the superfix r denotes rural areas) reads

$$\begin{aligned} \omega_r^r &= \psi \left(\theta H_r - l_r + \ln \left[\int_0^{\delta^{-1}} 2\pi s \left(1 - \delta s\right) \mathrm{d}s \right] \right) \\ &= \psi \left(\theta H_r - l_r + \ln \pi - 2 \ln \delta - \ln 3 \right). \end{aligned} \tag{11}$$

see equation (10). In the first line, $2\pi s$ is the circumference of the circle at distance s of the own location, $1 - \delta s$ is the fraction of the spillovers that survives at this distance, and $-l_r$ is the log population density (the inverse of average land use). By equation (6), average land use l_r depends on the price land and on average income, which themselves depend on ω_r, see equation (5): knowledge spillovers drive up the price of land, which in turn increases the population density and thereby the magnitude of the spill overs, yielding a self reinforcing loop. Substitution of these relations in equation (11) and solving for ω_r^r yields

$$\begin{aligned}\omega_r^r &= \Psi\left(-\overline{\theta}H_r + \lambda_\eta\lambda_z\alpha_T T_r + \ln\pi - 2\ln\delta - \ln 3\right), \\ \Psi &\equiv \frac{\psi}{1-\lambda_\omega\psi}.\end{aligned} \tag{12}$$

The parameter Ψ has to be positive for a bounded solution to exist. A bounded equilibrium therefore requires

$$\lambda_\omega\psi < 1. \tag{13}$$

If this condition were not satisfied, the reinforcing loop discussed before would explode. By equation (6) and (9), λ_ω is increasing in the substitution elasticity between land and other consumption, η. Were inequality (13) violated, either because the agglomeration parameter ψ were high or because people could easily adjust their land use (a high η), all economic activity would agglomerate at a single point in space with an infinite land price and all other land would be empty. For example, for $\eta = 1$ (Cobb Douglas utility) and $\alpha_z = 0$ (no residential agglomeration externalities), $\omega_\lambda = \overline{\lambda}/\lambda > 1$. Equation (13) would be easily violated in that case.

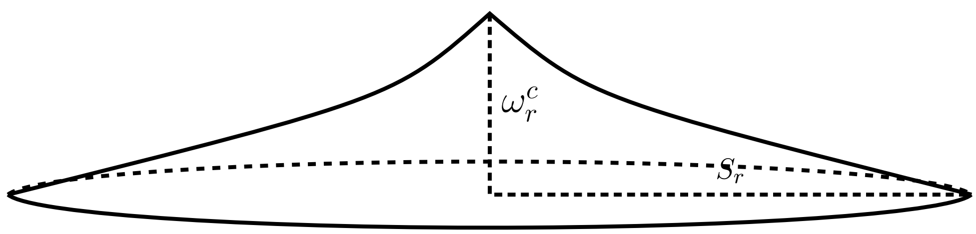
$$\eta = 1, \alpha_z = 0, \lambda_\eta = 1, \lambda_z = (\lambda - \alpha_z)^{-1}, \lambda_\omega \equiv \lambda^{-1}\left(\overline{\lambda} + \lambda_z\alpha_z\right)$$

The elasticity of the knowledge spillover ω_r^r with respect to regional mean level of education H_r is $-\Psi\overline{\theta} = \Psi(\theta - 1)$, which is positive for $\theta > 1$: a higher mean education level H_r increase spillovers. Spillovers raise house prices and therefore reduce land use, which allows a further increase in spillovers by a higher population density.

Where the labour supply curve dictates that wages should be decreasing in January temperature keeping land prices constant, see equation (5), the sign of this effect is reversed in the general equilibrium outcome. A high January temperature increases wages unequivocally. The reason is that its initial negative effect on wages is fully offset by higher land prices. In the absence of agglomeration spill overs, wages would be fully determined by labour demand and the positive effect of January temperature on utility would be offset by higher land prices. Hence, wages would be independent of land prices. However, as soon as agglomeration benefits come into play, a high January temperature raises the price of land relative to the price of other consumption. Workers substitute away from land to other consumption, thereby increasing the population density. This pushes up the agglomeration benefits. Hence, the labour demand curve dictates that wages are increasing in January temperature. The labour supply curve is satisfied by land prices rising by even more than the increase that is required to offset the positive supply effect of the amenity.

2.4.2 Cities

Like Lucas and Rossi-Hansberg (2002) and Rossi-Hansberg and Wright (2007), cities are assumed to have a circular shape. Like Rossi-Hansberg and Wright (2007), all employment is localized in the CBD at the city-center, which does not use any land at all; all jobs are therefore concentrated at a single point in space and hence, there is no loss in the transmission of ideas. Workers live in the area around the CBD. The cost of commuting to the CBD is a share κ or labour supply per unit of distance. Hence, somebody living at a distance s from the CBD works only a fraction $1 - \kappa s$ of her time; the remainder is lost during commuting. The worker's output and the knowledge spillovers she creates are similarly affected. Let S_r denote the edge of the city (the maximum distance from the CBD). By construction, $S_r < \kappa^{-1}$: commuting beyond that distance is useless as it leaves no time for working.



In Lucas and Rossi-Hansberg (2002) and Ahlfeldt et.al. (2015), land use varies within a city, depending on the price of land at a each location within the city. This is a more difficult structure than we are willing to handle for our purpose. We therefore introduce the artificial device of a city council that equalizes the private cost of commuting as a share of the wage rate $w_r(h)$ of the inhabitants living within city boundary S_r by imposing a balance budget system of taxes and subsidies. Inhabitants living close to the CBD pay a tax, while those living at the edge of the city receive a subsidy. These taxes and subsidies balance all commuting cost differentials within the city. The merit of this assumption is that it makes each location within a city equally attractive. Hence, land prices and the consumption of land are flat within the city.[4]

The city council sets the boundary S_r such that a worker living just outside the city is indifferent between commuting to the CBD to benefit from its agglomeration externalities and working at her home location, where she does not enjoy any agglomeration benefit at all.[5] Hence, ω_r^c (the

[4]This assumption is different from the concept of a city planner used in Rossi-Hansberg and Wright (2007) where cities compete for workers and where the social planner uses all land rents to compete for workers from other cities, thereby implementing a first best outcome. Our assumption only redistributes land rents and travell cost between the inhabitants of the city, but leaves the total amount of land rents unaffected.

[5]This assumption is not fully appropriate. Somebody living just outside the city benefits from the agglomeration benefits that apply in rural areas, see the previous section. However, since the nearby workers living inside the city work in the CBD, they do not contribute to the agglomeration benefits of people living just outside the city. Hence, somebody living just outside the city benefits from approximately half the agglomeration benefits that are enjoyed by somebody living far away from the city. For the sake of transparency, we choose to ignore these issues.

superfix c denotes cities) satisfies

$$\begin{aligned} \omega_r^c &= -\ln(1-\kappa S_r) \Rightarrow \\ \kappa S_r &= 1 - e^{-\omega_r^c}. \end{aligned} \tag{14}$$

Let n_r^c denote the city's log population and let f_r denote the log average net labour supply per worker after the deduction of commuting cost. Then

$$\begin{aligned} m_r &= n_r^c + f_r \\ f_r &= \ln\left[\int_0^{S_r} 2\pi s(1-\kappa s)\,\mathrm{d}s\right] - \ln\left[\int_0^{S_r} 2\pi s \mathrm{d}s\right] \\ &= \ln\left(1 - \frac{2}{3}\kappa S_r\right) = \ln\left(1 + 2e^{-\omega_r^c}\right) - \ln 3 \end{aligned} \tag{15}$$

Where ideas travel in rural areas, so that people don't have to commute and κ is therefore irrelevant, people commute in cities, so that ideas don't have to travel and δ is irrelevant. Define Δ to be

$$\Delta \equiv \ln\frac{\delta}{\kappa}. \tag{16}$$

Ahlfeldt et.al. (2015) report $\delta > \kappa$ and hence $\Delta > 0$, see their Table V. In the Appendix, we show that ω_r^c satisfies

$$\omega_r^c = \Gamma\left[\omega_r^r + 2\Psi\Delta\right], \tag{17}$$

where $\Gamma[\cdot]$ is differentiable function with

$$\Gamma'[\omega] \equiv \gamma[\omega] > 1,$$

see the Appendix for the derivation; γ is defined as $\gamma\left[\overline{\omega}_{r\in \mathrm{C}}^r + 2\Psi\Delta\right]$, where $\overline{\omega}_{r\in \mathrm{C}}^r$ is mean value of ω_r^r among city regions, $r \in \mathrm{C}$. The log population of a city n_r^c follows from inverting equation (10) and substitution of equation (15) for m_r

$$n_r^c = \psi^{-1}\omega_r^c - \theta H_r - \ln\left(1 + 2e^{-\omega_r^c}\right). \tag{18}$$

Let $n_H^c > 0$ and $n_{HH}^c < 0$ be the first and second partial derivatives of n_r^c with respect to H_r evaluated at the mean of $\overline{\omega}_{r\in \mathrm{C}}^c \equiv \Gamma\left[\overline{\omega}_{r\in \mathrm{C}}^r + 2\Psi\Delta\right]$, see the Appendix for details. The function $\Gamma[\omega]$ is depicted in Figure 1, joint with the regression lines and empirical observations for ω_r^r and ω_r^c for the data for 47 rural regions and 34 metropolitan areas in the US, see Section 3 below. The shape of the function justifies the following claims:

1. There is a critical threshold ω^o for ω_r^r below which an equilibrium of the urban form does

not exist.

2. Cities can only be a land price maximizing spatial structure (that is equivalent to: $\omega_r^c > \omega_r^r$, see equation (9)) if the cost of commuting κ are smaller than the spatial decay of knowledge spillovers δ (or: $\Delta > 0$). If not, it is always cheaper to let ideas rather than workers travel, since commuting reduces the time available for knowledge sharing *and* production, while the travel of ideas reduces only the former.[6]

3. For a sufficiently high value of Δ, there is a critical threshold ω^*, such that if $\omega_r^r < \omega^*$ the rural form is more efficient ($\omega_r^r > \omega_r^c$) while if $\omega_r^r > \omega^*$ the urban form is more efficient ($\omega_r^r < \omega_r^c$); $\omega^* \geq \omega^o$.

4. Since $\gamma > 1$, log wages ω_r are more sensitive to the average level of human capital and January temperature in cities than in rural areas. In the latter, the radius of the area around a location that contributes to the agglomeration benefits is fixed at δ^{-1}. In urban areas, the radius is determined endogenously, by equation (14). Hence, an increase in H_r raises not only the spillovers generated within a city of a fixed size; it also increases the size of the city.

5. Though the radius of a city is always increasing in H_r, this is not necessarily true for its population. Workers with high human capital earn a higher income and therefore consume more land. The effect of H_r on the city's radius can be offset by its effect on the population density. The larger ω_r^c, the more likely it is that an increase in human capital requirements reduces the city's population.

Since ω_r^r is increasing in the average level of human capital and in January temperature, regions with a high demand for human capital H_r and an agreeable January temperature generate a higher land price when organized as a city than as a rural area. A high January temperature increases land prices and therefore leads to a higher population density, which is conducive to agglomeration benefits. For $\theta > 1$, this makes a city more attractive for the production of tradeables with a high demand for human capital H_r. There is therefore a good reason for the IT industry to be located in California: the agreeable climate supports high house prices, which yields the high density that is conducive to the knowledge spill overs desired by the IT industry.

2.5 Equilibrium

An equilibrium to this multi-region economy is a set of fixed log wage effects ω_r, log average land use l_r, and log land prices p_r (for all regions), and the log population size n_r^c (for cities) as a function of mean level human capital H_r, the January temperature T_r, and their spatial organization (urban

[6] This can be see immediately by realizing that equation (14) implies that $S_r < \kappa^{-1}$. If $\kappa > \delta$, then the radius of area for which a rural region extracts knowledge spill-overs exceeds that for which a city can do.

Figure 1: $\Gamma(\omega)$ and Empirical Estimates of the Wage Fixed Effect

Note: The plot shows $\omega_r^c = \Gamma(2\Psi^r\Delta + \omega_r^r)$ and $\omega_r^r = \Psi^r\left(-\bar{\theta}H_r + \lambda_\eta\lambda_z\alpha_T T_r\right)$, with model parameter estimates in Table 3. The red curve is the city agglomeration effect ω_r^c of area r. The blue line ω_r^r is the non-city agglomeration effect of area r.

versus rural). These variables must satisfy equations (9), (12), (17), and (18). Since we have data on house prices - not on land prices - we focus on the value of the land of a representative house in region r, in logs: $v_r \equiv l_r + p_r$. The full model reads

$$\begin{aligned}
\omega_r^r &= \Psi\left(-\bar{\theta}H_r + \lambda_\eta\lambda_z\alpha_T T_r\right), \qquad (19)\\
\omega_r^c &= \Psi\cdot\Gamma\left[-\bar{\theta}H_r + \lambda_\eta\lambda_z\alpha_T T_r\right] \cong \gamma\Psi\left(-\bar{\theta}H_r + \lambda_\eta\lambda_z\alpha_T T_r\right),\\
v_r &= (\lambda_z - \lambda_\omega)\,\omega_r + H_r + \overline{\lambda\eta}\lambda_z\alpha_T T_r,\\
n_r^c &= \psi^{-1}\omega_r^c - \theta H_r - \ln\left(1 + 2e^{-\omega_r^c}\right), \quad \frac{\partial n_r^c}{\partial\omega_r^c} > 0; \quad \frac{\partial^2 n_r^c}{(\partial\omega_r^c)^2} < 0,
\end{aligned}$$

where we linearize equation (17) in the second equality in the second line and where $\lambda_\eta, \lambda_\omega, \lambda_z$, and Ψ are defined in equation (6), (8), (9), and (12). The relation between the log population of city and and the fixed effect in wages is concave. The closer the radius of a city comes to the physical constraint imposed by the commuting cost, $S_r = \kappa^{-1}$, compare equation (14), the smaller are the opportunities for further extension of the city. The population might eventually decline, when the radius is close to its maximum κ^{-1}: a further increase in ω_r will then increase the benchmark utility level H_r of the median citizen. High wealth individuals can afford large lot sizes, thereby reducing the city's population.

These equations can be estimated, using data on ω_r, v_r, H_r and T_r for both rural regions and cities and on n_r^c for cities. We apply estimates on the parameters λ, η, ψ, and α_z from other sources, see Section 3.4, which yields expressions for the composite parameters $\lambda_\eta, \lambda_z, \lambda_\omega$ and Ψ. Then, the estimation of equation (19) yields a number of testable predictions:

$$\begin{aligned} \gamma &= \frac{\omega_H^c}{\omega_H^r} = \frac{\omega_T^c}{\omega_T^r} = \frac{v_H^c - 1}{v_H^r - 1} > 1, \\ \lambda_z - \lambda_\omega &= \frac{v_H^c - 1}{\omega_H^c} = \frac{v_H^r - 1}{\omega_H^r} \\ \lambda_z \alpha_T &= \frac{\omega_T^c}{\gamma \Psi \lambda_\eta} = \frac{v_T^c - (\lambda_z - \lambda_\omega)\, \omega_T^c}{\overline{\lambda \eta}} \\ \theta &= 1 + \frac{\omega_H^r}{\Psi} = 1 + \frac{\omega_H^c}{\gamma \Psi} \end{aligned} \tag{20}$$

where the subscripts H and T on a variable denotes the partial derivatives of that variable with respect H_r and T_r respectively. Since these functions are approximately linear, these partial derivatives do not depend on r. Hence we omit the subscript r. The first line of equations (20) shows three independent ways to estimate γ, while the second line shows two independents ways to estimate $\lambda_z - \lambda_\omega$. Multiple estimates of a single parameter provide over-identifying restrictions and hence tests of the model. Subsequent lines provides ways to estimate $\lambda_z \alpha_T$ and θ.[7] In the next section, we use sign restriction on the partial derivatives and the over-identifying restrictions (20) to test the model.

3 Empirical evidence

3.1 Data

For the estimation of the model, we need data on wages, human capital, housing values, and January temperature for a set of regions. We draw data from four different sources. Individual level data are taken from the Current Population Survey, Merged Outgoing Rotation Groups (CPS-MORG) from 1979 till 2015. We use the hourly wage, years of education, occupation, industry and other demography information as gender, age, marital status, and race. Our sample includes all workers age 16 to 64.

For our classifications of regions, we select 34 MSAs for which both individual level and regional level data are available. We then take the remaining part of each state as one non-city

[7] The two ways to estimate θ do not provide an additional overidentifying restriction, as the information on ω_H^r and ω_H^c is already used in the first line for the estimation γ.

region. From 1979 to 1985, we use 1970 Census ranking to identify MSAs; from 1986 to 1988, we use the CMSA and PMSA identifier; from 1989 to 2003, we use MSAFIPS, for the remaining years, we use CBSAFIPS. We exclude Hawaii and Alaska. Furthermore, New Jersey drops out since all of its area is part of NY-NJ MSA, leaving us with 34 MSAs and 47 rural areas.

The regional population and employment data are taken from US Census Bureau. The Housing Price Index (HPI) is taken from the Federal Housing Finance Agency All-Transaction Indexes, both for MSAs and for the non-MSA part of states. To make the housing price comparable across regions, we also calculated the housing value index, using the additional information from Zillow Research, Zillow Home Value Index. We use the estimated median home value for all homes within a region. State level data for Maine and Louisiana are missing from this data set. Instead, we use the average home value at the county level. Average January temperature data is from US National Oceanic and Atmospheric Administration, 1981-2010 US climate normals, see Glaeser (2004). All data cover the period from 1979 to 2015 at annual frequency.

3.2 The construction of the human capital index h

Estimation of the model requires data on the average level of human capital H_r and on the fixed effect ω_r in each region. Measuring both variables is a challenge since human capital is only partially observed. The observed part of human capital $\widehat{H}_r$ in a region is likely to be a biased estimate of the actual level of human capital H_r, since the observed and the unobserved part of human capital are likely to be positively correlated. Then, the interregional variation in the observed part is an underestimation of the actual variation in the level of human capital. If so, the regional fixed effect ω_r picks up this effect unobserved human capital. Hence, its estimate is upwardly biased in regions with high H_r. We propose a simple correction for this problem based on this idea of substitutability of observed and unobserved human capital and we provide empirical evidence supporting our method.

An index for the observed human capital of individual i is constructed by means of a simple log linear wage equation[8]

$$\begin{aligned} w_i &= \omega_r + h_i + e_i, \qquad (21) \\ h_i &\equiv \widehat{h}_i + \underline{h}_i, \\ \widehat{h}_i &\equiv \omega' x_i, \end{aligned}$$

where w_i is the observed log hourly wage for worker i working in region r, where $\widehat{h}_i$ and $\underline{h}_i$ are the

[8]The linearity of this equation is not an important restriction on the generality of the analysis, see Gautier and Teulings (2008). Suppose wages are an increasing but non-linear function of some human capital index h^*: $w = w(h^*) = w(\omega' x)$, with $w'(h^*) > 0$. By defining a transformed human capital index $h = w(h^*)$ the linearity can be imposed without loss of generality. The non-linearity in the relation with the vector x can then be addressed by applying a polynomial in x.

observed and unobserved part of human capital respectively, where e_i is measurement error in the observed log wage, and where x_i is a vector of standard personal characteristics like age, years of education, gender, marital status, and race. Without loss of generality, we assume the components $\widehat{h}_i$ and $\underline{h}_i$ to be orthogonal over the full sample of all regions. For the sake of convenience the index r denotes both the region and the moment of time at which it is observed; hence, there are $\mathrm{R} \times \mathrm{T}$ values of ω_r, where $\mathrm{R} = 81$ denotes the number of regions and $\mathrm{T} = 37$ denotes the number of years. Since we use cross-section data, an individual is observed only in one region and at one point in time, so i implies r. The parameter ω_r is a fixed effect for region r, and ω is a vector of parameters of the same dimension as the vector x_i which is common to all regions; the parameter vector aggregates the components x_i into the single human capital index $\widehat{h}_i$. The time dummies ω_t capture nation-wide inflation and productivity growth. In practice, we estimate equation (21) by OLS without these time dummies and then demean the estimated ω_r for each time t. Referring to $\widehat{h}_i$ as the observed human capital of a worker glosses over all kind of hairy issues like whether the effect of gender or race might attributed to differences in human capital or that these variables might be proxies for all kind of other processes, like interrupted careers of women or discrimination against women and blacks. Since our aim is just to agglomerate all observable variables in a single index that reflects the earning capacity of a worker, we sidestep these issues.

For the sake of convenience, w_i and all elements of x_i are demeaned over the full sample. Hence

$$\mathrm{E}\left[w_i\right] = \mathrm{E}\left[\widehat{h}_i\right] = \mathrm{E}\left[\underline{h}_i\right] = \mathrm{E}\left[\omega_r|t\right] = 0. \tag{22}$$

Though the overall means of $\widehat{h}_i$ and $\underline{h}_i$ are zero, their mean within an individual region might be different from zero due to selective migration. We define the mean level of observed human capital in region r as

$$\widehat{H}_r \equiv \mathrm{E}\left[\widehat{h}_i|r\right].$$

H_r is defined similarly as the mean of h_i for all workers in region r; the means of $\widehat{H}_r$ and H_r over the full sample are zero by construction (since x_i is demeaned over de full sample and $\mathrm{E}[\underline{h}_i] = 0$).We define

$$\omega_r = \widehat{\omega}_r + \mathrm{E}\left[\underline{h}_i|r\right]. \tag{23}$$

Estimation of equation (21) by standard techniques yields an estimate of $\widehat{\omega}_r$, but not of ω_r, because $\underline{h}_i$ and hence $\mathrm{E}[\underline{h}_i|r]$ is unobserved. The Proportionality Assumption below fills this gap.

When we want to use observed human capital $\widehat{H}_r$ as an estimate total human capital, we have to apply a correction for effect of unobserved human capital. For this purpose, we use the single index assumption, see Teulings (1995): the earning capacity of a worker can be meaningfully summarized in a single index $h_i = \widehat{h}_i + \underline{h}_i$, as in equation (21). The critical assumption here is not the unit coefficient on both components $\widehat{h}_i$ and $\underline{h}_i$[9] or by the linearity of the components[10], but in the additivity, implying that observed and unobserved human capital are perfect substitutes: each unit of observed human capital can be replaced by a unit of unobserved human capital at fixed rate of transformation. As a logical implication of this feature, one would expect that a region that selects workers with high human capital h_i selects them both along the observed and the unobserved dimension. This leads to the Proportionality Assumption stated below.

The Proportionality Assumption

When a worker has human capital $h_i = \widehat{h}_i + \underline{h}_i$, then the index h_i is a sufficient statistic for the expectation of its observed and unobserved component. In particular, the following relations apply

$$\begin{aligned} \mathrm{E}\left[\widehat{h}_i | h_i, r\right] &= R^2 H_r \\ \mathrm{E}\left[\underline{h}_i | h_i, r\right] &= \left(1 - R^2\right) H_r, \end{aligned} \tag{24}$$

where

$$R_h^2 \equiv \frac{\mathrm{Var}\left[\widehat{h}_i\right]}{\mathrm{Var}\left[\widehat{h}_i\right] + \mathrm{Var}\left[\underline{h}_i\right]}. \tag{25}$$

where $\mathrm{Var}\left[\widehat{h}_i\right]$ and $\mathrm{Var}[\underline{h}_i]$ are the variance of $\widehat{h}_i$ and $\underline{h}_i$ in the nation-wide populations.

This assumption is a natural extension of the idea that observed and unobserved human capital are perfect substitutes and that the decomposition of h_i in both components is irrelevant for the earning capacity of the worker. Taking expectations for region r in equation (24) yields a simple expression for H_r and ω_r as a function of $\widehat{H}_r$ and $\widehat{\omega}_r$:

$$\begin{aligned} H_r &= R_h^{-2} \widehat{H}_r, \\ \mathrm{E}\left[\underline{h}_i | r\right] &= \frac{1 - R_h^2}{R_h^2} \widehat{H}_r, \\ \omega_r &= \widehat{\omega}_r - \frac{1 - R_h^2}{R_h^2} \widehat{H}_r. \end{aligned} \tag{26}$$

We use equation (21) to estimate the parameter vector ω, the region fixed effects $\widehat{\omega}_r$, and the observed human capital, both per individual $\widehat{h}_i$ and the regional mean $\widehat{H}_r$. The exact composition of the vector x_i and estimation results for the parameter vector ω are presented in the Appendix. Three examples of the index $\widehat{h}_i$ characterize the distribution. The 10th percentile of the distribution of $\widehat{h}_i$ is -0.426; a typical worker in this group is a black married female with 10 years of education

[9] Deviations could be eliminated by a simple redefinition of the unit of measurement of h_i

[10] Consider an alternative index $\widehat{h}_i^*$ that enters as $h_i = h\left(\widehat{h}_i^*\right) + \underline{h}_i$ where $h(\cdot)$ is an increasing function, then we can replace $h\left(\widehat{h}_i^*\right)$ by $\widehat{h}_i = h\left(\widehat{h}_i^*\right)$

Table 1: Summary Statistics

Individual level		Var Decomposition %			
Variable	S.D.	Time	Region	Time x Region	Residual
w_{ir}	0.6817	30.3	2.6	0.3	66.8
$\widehat{h}_i$	0.3498	5.6	0.6	0.2	93.6
Regional level		Correlation Matrix			
Variable	S.D.	$\widehat{H}_r$		$\widehat{\omega}_r$	
$\widehat{H}_r$	0.0281	1			
$\widehat{\omega}_r$	0.0836	0.6457		1	

and 26 years of experience. The median value of $\widehat{h}_i$ is -0.012, corresponding to a white married male with 12 years of education and 8 years of experience. The 90 percentile is 0.429, which corresponds to a white married female with 18 years of education and 21 year of experience. Clearly, the median level of $\widehat{h}_i$ should be close to zero by construction, since we demeaned x_i across the full sample; the only reason for the medians not being exactly zero is that the median is not equal to the mean.

The estimation results for $\widehat{\omega}_r$ and $\widehat{h}_i$ are summarized in Table A2. Most of the variance in observed human capital is within regions; only 0.6% of the variance is between regions. Since the private return to human capital is normalized to unity, observed human capital explains only only $(0.0281/0.0836)^2 = 11\%$ of the interregional variation in wages.

The intra-regional variance in log wages can be decomposed into three orthogonal components: observed human capital, unobserved human capital, and measurement error in wages. Let E be the share of the measurement error in the variance of observed log wages. We use an independent estimate of $E = 0.30$ by Angrist and Krueger (1999). Since we observe the variance of observed human capital, the variance of the unobserved human capital can be backed out as a residual item. Applying the Proportionality Assumption, see equation (26), we obtain

$$
\begin{aligned}
R_h^2 &= 0.53, \\
H_r &= 1.89\widehat{H}_r, \\
\omega_r &= \widehat{\omega}_r - 0.89\widehat{H}_r.
\end{aligned}
$$

Applying this correction to the interregional variance in observed human capital, human capital explains $1.89^2 \times 11\% = 39\%$ rather than 11% of the interregional variation in log wages.

This calculation depends critically on the Proportionality Assumption. There are several ways to put this assumption to an empirical test. Single Index Models are a special case of Rosen's (1974) hedonic pricing model, see e.g. Sattinger (1975), Teulings (1995, 2005), and Gabaix and Landier (2007). In these models, workers with high human capital have a comparative advantage in complex jobs. In a market equilibrium, they are therefore assigned to such jobs. Since wages are increasing in human capital, workers in complex jobs have both more human capital and higher wages. Hence, we can use a measure of job complexity as an instrument for human capital. We use 3 digit occupational dummies. As a first stage regression, we regress w_i on the occupation dummies and construct an occupation complexity index $o_i = \beta' d_i$ as the explained part of this regression. Let R_o^2 be the R^2 of this regression, after deducting measurement error in w_i; R_o^2 happens to be exactly equal to R_h^2: 0.53. This index can serve as an instrument for both the observed and the unobserved part of human capital. Moreover, we can use log wages minus the part explained by observed human capital as a proxy for unobserved human capital: $w_i - \widehat{h}_i = \underline{h}_i + \varepsilon_i$ (unobserved human capital plus measurement error in wages, see equation (21)). We therefore regress both $\widehat{h}_i$ (observed human capital) and $w_i - \widehat{h}_i$ (unobserved human capital) on the occupational complexity index o_i. If the Proportionality Assumption were perfect the R^2 of the regression using $\widehat{h}_i$ should be $R_h^2 R_o^2 = 0.28$, while the R^2 of the regression $w_i - \widehat{h}_i$ should be $(1 - E)\left(1 - R_h^2\right) R_o^2 = 0.17$. The actual R^2s are 0.36 and 0.12 respectively, so not exactly in line with the prediction, but occupational complexity clearly correlates with our proxy for unobserved human capital.

We can extend this idea one step further, by estimating separate β-vectors for $\widehat{h}_i$ and $w_i - \widehat{h}_i$. If $\widehat{h}_i$ and $\underline{h}_i$ measure different aspects of human capital which are not perfect substitutes, as in a Double Index Model e.g. with both intellectual ability and social intelligence as inputs, they span a two– rather than an one-dimensional space. Different combinations of $\widehat{h}_i$ and $\underline{h}_i$ make a worker apt for different occupations even when their sum h_i is the same. Since our occupational classifications has more than 300 entries, it can be expected to span this two dimensional space. The linear combination that correlates best to the observed component $\widehat{h}_i$ should therefore be different from the linear combination that correlates best to the unobserved component $\underline{h}_i$. The β-vectors for both regressions (for $\widehat{h}_i$ and $w_i - \widehat{h}_i$ respectively) and o_i-indexes derived from them are the same, the Single Index Model applies. The actual correlation between both o_i-indexes is 0.70. Hence, the Single Index Model gives a fairly accurate though not perfect description of the data. In the IV interpretation of using occupation data as an instrument for human capital, the correlation between both o_i-indexes is equivalent to the standard over-identification test of the instruments. If the correlation coefficient were equal to unity, the test would be accepted. As a point estimate, the null is clearly rejected. However, the high correlation coefficient shows the Single Index Model to be a reasonable first order approximation.

The correction of H_r and ω_r for unobserved human capital can be considered as the limiting case, where interregional variation observed and unobserved human capital, $\widehat{H}_r$ and $\underline{H}_r$ is perfectly correlated. The inferred interregional variation of H_r can therefore be viewed as an upper-bound for the true variation. Since this variation enters negatively in the corrected estimate for the regional fixed effect in log wages ω_r, see equation (26), this corrected estimate is a lower bound for the true variation in ω_r; this estimate can therefore viewed as lower-bound for the true human capital externalities.

3.3 Estimation

In our model, average human capital and January temperature are treated as exogenous regional characteristics. While most economists would be willing to buy the exogeneity of January temperature, they are unwilling to accept to exogeneity of the average human capital in a region. One would expect that the level of human capital in a region responds to all kind of changes, like better amenities, a greater supply of houses, and a greater availability of suitable jobs for people with a particular level of human capital. Which of these factors is the ultimate driver of changes in human capital is not a prior clear. We need therefore an instrument for the exogenous variation in the regional level of human capital. The regional industrial structure has a long memory, see Amior and Manning (2018). Differences in industrial structure can therefore be helpful to explain the interregional differences in average human capital and their evolution over time. We construct two versions of this Bartik instruments for the regional average level of human capital, see Goldsmith-Pinkham, Sorkin, and Swift (2018) for a discussion.

The first instrument weighs nation-wide changes in the industry mix and the average level of education in each industry (both excluding the own region) by the regional average industry mix

$$\widetilde{H}^{B}_{rt} = \sum_{j} \frac{E_{rj}}{E_r} g_{(-r)jt} H_{(-r)j}$$

where H_j is the nation-wide mean level of human capital in industry j (where we use the subscript $(-r)$ to denote that we exclude region r from the calculation of the nation wide mean), where E_{rjt} is the employment in region r in industry j at time t, and where $g_{(-r)tj}$ is the nation-wide evolution of employment in industry j relative to the evolution of total employment $E_{(-r)t}$, $g_{(-r)jt} = \frac{E_{(-r)jt}}{E_{(-r)j}} - \frac{E_{(-r)t}}{E_{(-r)}}$. This instruments put the weight on nation-wide changes in the industry mix. If a region has a larger share of growing industries with a high human capital requirement, the demand for human capital in the region goes up. We refer to this instrument as the between-industry instrument. Since $\mathrm{T}^{-1}\Sigma_{t=1}^{\mathrm{T}} E_{(-r)jt} = E_{(-r)j}$, $\Sigma_{t=1}^{\mathrm{T}} g_{(-r)jt} = 0$, this instrument is undefined for the cross-section.

The second instrument weighs nation-wide changes in the human capital requirement within industries by the regional average industry mix

$$\widetilde{H}^{W}_{rt} = \sum_{j} \frac{E_{rj}}{E_r} H_{(-r)jt}$$

If a region has a larger share of industries with an increasing human capital requirement, the demand for human capital goes up. We refer to this instrument as the within-industry instrument.

furthermore, $T^{-1}\Sigma_{t=1}^{T} H_{(-r)jt} = H_{(-r)j}$. This instrument exists for both the panel and the cross-section version of the model.

Table 2 presents the estimation results, both for the cross-section (left panel) and the time-series evidence (right panel). Column (1) gives the results for the first stage regression for the cross-section and the time-series part of the analysis. We enter both the instrument and its square. In both cases, the instrument is strong. In the cross-section, the squared instrument is significant. This might capture an agglomeration effect in innovation intensive industries, compare Desmet and Rossi-Hansberg (2009): if a region has an above average share in an industry, then it is likely to be the innovation hub of that industry. Hence, an increase in the human capital requirement of that industry will disproportionally favour this region. Column (2) and (3) uses the instrument to analyze the cross-section effect of education on housing values and the wage fixed effect. Using the instrument for human capital renders the city dummy insignificant (city-size is not included, as it is endogenous in the model). January temperature and human capital have the predicted sign. The positive effect of January temperature on house prices is a standard prediction: house prices would pick up the amenity-value. However, in a model without agglomeration benefits, one would expect wages to be independent of January temperature: wages would be determined by nation-wide product market competition. In a model with agglomeration effects, the upward effect January temperature on house prices reduces average lot-size and therefore increases the agglomeration benefits, pushing up wages. Our normalization of the private return to human capital to unity provides an easy interpretation of the coefficient on human capital in the regressions for ω_r: for example, a coefficient of 0.912 in the cross-section regression implies that if the level of human capital of all workers in a region is increased by one unit, this would yield a private return of unity and an additional "social" return from agglomeration benefits of 0.912. These numbers square well with the high social return to human capital reported by Genniola et.al. (2013). Splitting the sample in cities and rural areas shows that the effect of human capital and January temperature on house prices and the wage fixed effect to be larger in cities than in rural areas as the model predicts. None of the variables is able to explain the variation in city-size, see column (4).

The time-series results for the within instument are in column (5) to (8), the results for the between instrument are in column (9) to (12). In both cases, the instruments are strong. Human capital has strong positive impact on house prices and the wage fixed effect, in the sample as a whole and in the sub-samples for cities and rural areas, where again the effect are more positive for cities than for rural areas. The results for the between and the within instrument are very similar, though the coefficients tend to be somewhat larger for the within instrument, in partilar for house prices. The Haussman overidentification test of the instruments is even rejected in some cases. Moreover, the results have similar orders of magnitude as the cross-section results. The results for city-size are surprisingly strong, see column (8): the first derivative is positive, while the second is negative, exactly as theory predicts. Both coefficients are highly significant.

3.4 Identification and testing

We apply the estimation results to obtain rough estimates of the parameters of the model. Moreover, we get some idea of the validity of the over-identifying restriction listed in equation (20). A number of standard parameters are derived from the literature. Doing so will tie our hands when trying to match our estimation results. Ahlfeldt et.al. (2015) assume a floor space share for residential use of 25% and for commercial use of 20%, adding up to 45%. Using a land share of 40% in the value of real estate, the land share λ in consumption is about 0.18. They report the agglomeration elasticity for consumption α_z to be 0.10 and for production ψ to be about 0.08, see their Table V. Albouy, Ehrlich and Liu (2016) and Teulings, Ossokina, and De Groot (2018) report evidence that the elasticity of substitution η is less than unity. We use a value of a half:

$$\alpha_z = 0.10, \qquad \psi = 0.08, \qquad \lambda = 0.18, \qquad \eta = 0.50.$$

Using these numbers, we obtain:

$$\begin{aligned} \lambda_\eta &\equiv 1 - \overline{\lambda}\overline{\eta} = 0.59, \\ \lambda_z &\equiv \left(\lambda - \lambda_\eta \alpha_z\right)^{-1} = 8.3, \\ \lambda_\omega &\equiv \lambda^{-1}\left(\overline{\lambda}\eta + \lambda_\eta^2 \lambda_z \alpha_z\right) = 3.9, \\ \Psi &\equiv \frac{\psi}{1 - \lambda_\omega \psi} = 0.12, \end{aligned} \tag{27}$$

Note the importance of allowing for residential agglomeration benefits. If α_z were equal to zero, λ_z would be 5.6 instead of 8.3, while λ_ω would be 2.3 instead of 3.9. Hence, $\lambda_z - \lambda_\omega$ would be 2.3 rather than 4.4, much lower than the value that we estimate, see below.

Table 3 provides a rough estimate for the remaining parameters. We just provide point estimates, based on the estimation results in Table 2. We don't report point estimates when either the numerator or the denominator has a t-statistic of less than 0.5. Using the formula for the standard deviation of the ratio of two independent stochastic variables, we conclude that non of the overidentifying restrictions is rejected. For the time-varying variable H_r, we have three estimates (cross-section and two instruments for the time-series). For the time-invariant variable T_r, we have only cross-section estimates. Altogether, we have ten different estimates for the excess sensitivity of house prices and wages in cities relative to rural areas γ, which is predicted to be larger than one. The t-statistics of either the numerator or the denominator of three of these estimates are below the 0.5 threshold. The mean of the remaining seven estimates is roughly $\gamma = 2$. Only the estimate based on $\left(v_H^c - 1\right) / \left(v_H^r - 1\right)$ deviates substantially, but this difference is not significant.

Table 2: Bartik IV Regression Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	H_r	v_r	w_r	n_r	H_{rt}	v_{rt}	w_{rt}	n_{rt}	H_{rt}	v_{rt}	w_{rt}	n_{rt}
Bartik IV	W	W	W	W	W	W	W	W	B	B	B	B
Time Dummy	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	Y
Panel A: City and Non-City Sample												
ln Jan Temp	-0.0504	0.420	0.0831									
	(-4.71)	(3.85)	(3.79)									
Bartik IV	1.074				1.162				0.972			
	(4.70)				(10.32)				(10.68)			
Bartik IV sq.	9.412				-0.056				-1.698			
	(2.20)				(-0.60)				(-2.53)			
Human Capital Index		5.093	0.912			5.783	-0.0722			2.131	0.663	
		(4.08)	(3.64)			(8.12)	(-0.70)			(4.78)	(5.12)	
Metro Dummy	0.0395	-0.158	0.0238									
	(4.36)	(-1.55)	(1.16)									
Observations	81	81	81		2,997	2,997	2,997		2,997	2,997	2,997	
Panel B: City Sample												
ln Jan Temp		0.786	0.0928	4.021								
		(4.15)	(2.50)	(0.75)								
Human Capital Index		7.138	1.186	71.89		10.61	0.735	4.910		5.144	0.542	4.046
		(3.99)	(3.38)	(0.69)		(4.74)	(2.73)	(3.20)		(4.55)	(2.60)	(4.29)
Human Capital Index sq.				-662.3				-3.886				-4.288
				(-0.66)				(-5.19)				(-5.26)
Observations		34	34	34		1,258	1,258	1,258		1,258	1,258	1,258
Panel C: Non-city Sample												
ln Jan Temp		0.0930	0.0633									
		(0.72)	(2.24)									
Human Capital Index		1.817	0.494			2.373	-0.0903			0.642	0.535	
		(1.08)	(1.35)			(5.18)	(-0.86)			(1.44)	(3.72)	
Observations		47	47			1,739	1,739			1,739	1,739	

Note: Robust t-statistics in parentheses.

We have eight independent estimates of the effect of human capital on the agglomeration force, θ. Ignoring the one case where the estimate is insignificant, all estimates are in the same ballpark. This ballpark is similar as reported by Gennaioli et.al. (2013), who report $\theta = 6$. The parameter $\lambda_z - \lambda_\omega$ can be derived from our a priori assumptions, see equation (27). Except for one estimate, all estimates fit this value. The estimate for the parameter $\lambda_z \alpha_T$ are all around the ballpark of 0.9.

The parameters $\Delta = \ln(\delta/\kappa)$ and ω_0, see equation (10), are derived by means of the following argument, see Figure 1. Let ω_r denote the actual value of the fixed effect in log wages in region r as estimated for the wage data and let $\widehat{\omega}_r^r$ be the predicted value of ω_r^r from the data on human capital H_r and January temperature T_r, using the estimation results in Table 2 for rural areas. Figure 1 plots the points ω_r as a function of $\widehat{\omega}_r$ for both rural and urban regions; red crosses and blue circles respectively. By construction, the red regression line for rural areas coincides with the main diagonal. The point on this line denotes the mean value $\overline{\omega}_{r\in \mathrm{C}}^r$ for cityies if they were organized as rural regions, $r \in \mathrm{C}$. The blue line is the regression line for cities. The point on this regression line corresponds to $\overline{\omega}_{r\in \mathrm{C}}^r$. The slope of the blue line is $\gamma = 2$. Let Δ_Γ be the gain in log wages of being organized as a city rather than as a rural region for $\omega_r = \overline{\omega}_{r\in \mathrm{C}}^r$. This wage gain is 7.5%. We identify Δ and $\overline{\omega}_r^c$ as the solution to the system

$$\begin{aligned} \gamma &= \gamma\left[\overline{\omega}_{r\in \mathrm{C}}^r\right], \\ \Delta_\Gamma &\equiv \Gamma\left[\overline{\omega}_{r\in \mathrm{C}}^r + 2\Psi\Delta\right] - \overline{\omega}_{r\in \mathrm{C}}^r. \end{aligned} \tag{28}$$

The solution is elaborated in the Appendix. The parameter ω_0 in equation (10) serves as a mean shifter that sets $\overline{\omega}_{r\in \mathrm{C}}^r$ to the appropriate value. We find $\Delta = 1.09$, which is lower than reported in Ahlfeldt et.al. (2015), who report $\Delta = \ln\delta/\kappa = 2.70$, see their Table 5. The figure shows that $\omega^* = \omega^o$: even for smallest value of ω_r^r for which is city is feasible, the urban form is more efficient than the rural form. Quite a number of rural regions would be better off by organizing themselves as cities as to benefit from the higher knowledge spill overs. Seen through the lens of our model, these regions have been unable to solve the coordination problem of where within the area of the city to locate their CBD. Hence, they got stuck in the rural form with dispersed economic activity. In this model where all investment in housing is reversible, a city can never get stuck in the city form while the rural form is more efficient: if ω_r would drop below $\omega^* = \omega^o$, the city would simply collapse.

Relative to the other parameters, which fit our theoretical expectations reasonalby close, the estimates for the elasticities n_H^c and n_{HH}^c do have the right sign but are a factor two to four higher than our theoretical prediction.

Although the model treats the mean educational requirement H_r of a region's activities as exogenously determined by long run forces governing the regional industrial structure, this structure will be endogenous in the very long run. Entrepreneurs setting up activities that require an

Table 3: Structural identification and over-identification restrictions

	cross-section	time-series	time-series
γ	2.0	rough mean estimate	
$\frac{\omega_H^c}{\omega_H^r}$	2.4	$-$	1.2^{**}
$\frac{\omega_T^c}{\omega_T^r}$	1.5^{*}	$n.a.$	$n.a.$
$\frac{v_H^c-1}{v_H^r-1}$	$-$	6.7^{**}	$-$
θ	5	rough mean estimate	
$1+\frac{\omega_H^c}{\Psi^c}$	6.1^{**}	4.2^{**}	3.7^{**}
$1+\frac{\omega_H^r}{\Psi}$	5.3	$-$	5.3^{**}
$\lambda_z-\lambda_\omega$	4.4	from equation (27)	
$\frac{v_H^c-1}{\omega_H^c}$	5.2^{**}	13.1^{**}	7.6^{**}
$\frac{v_H^r-1}{\omega_H^r}$	1.7	$-$	$-$
$\lambda_z\alpha_T$	0.85	rough mean estimate	
$\frac{\omega_T^c}{\gamma\Psi\lambda_\eta}$	0.7^{**}	$n.a.$	$n.a.$
$\frac{\omega_T^r}{\Psi\lambda_\eta}$	0.9^{*}	$n.a.$	$n.a.$
$\frac{v_T^c-(\lambda_z-\lambda_\omega)\omega_T^c}{\bar{\lambda}\bar{\eta}}$	0.9^{**}	$n.a.$	$n.a.$
n_H^c	7.4	from Appendix, see equation (31)	
n_H^c	$-$	4.9^{**}	4.3^{**}
n_{HH}^c	-18.2	from Appendix, see equation (31)	
n_{HH}^c	$-$	-4.3^{**}	-3.9^{**}

Note: n.a.: not available; -: lowest t-stat<0.5; blank : lowest t-stat<1.5; *: lowest t-stat<2.5; **: lowest t-stat>2.5.

extensive exchange of ideas will benefit from the existing urban infrastructure. Boston's supply of well educated workers attracted research intensive firms. The model predicts that cities have a comparative advantage in human capital intensive activities. Table A3 list the 87 regions in our sample by their mean level of human capital over the sample period. Cities rank high in this table, as the model predicts. The case of Riverside/Los Angeles is interesting. Though the city opened one of the first underground lines in the US in 1925, this line was closed down in 1955. The city became to rely heavily on car transport, which is land-intensive and therefore does not allow the high densities that are conducive to agglomeration benefits (a low commuting cost κ). This might explain why the city is less attractive to activities with a high human capital requirement.[11]

4 Counterfactuals

We use the model for the calculation of several counterfactuals. We use the linearized version of the model in equation (19) with full version for the function $\Gamma[\cdot]$ and equation (7) for the log average land use. First, we calculate a reference point that is consistent with the model, using the data on H_r and T_r; we use the average value of H_r for the period 1977-2015. We to calculate ω_r^r for all regions. Then, we solve the implicit function $\Gamma[\omega_r^r + 2\Psi\Delta]$ to obtain values for ω_r^c and n_r^c for cities. These data are use to calculate the log average land use l_r for all regions. For rural regions, we use the actual population n_r^r as as a point of reference. For cities, we use the calculated population n_r^c. The total land available for economic use, A_r, is then equal to

$$A_r = \exp(n_r + l_r).$$

We keep this endowment of land for each region fixed in all counterfactuals. The radius S_r of cities is determined endogenously and hence the land area A_r. However, since there are no cities is both counterfactuals, this does not affect the calculation of the counterfactual. The counterfactual keeps the land area constant and calculates the counterfactual population by dividing this land area by the counterfactual average land use.

We consider two counterfactuals. In the first, we rule out the city form. All regions have to take the rural form. In the second, we assume that there are no agglomeration externalities on the production side, so $\psi = 0$, although we do allow for agglomeration benefits on the consumption side, via the log income per unit of land, z_r. We calculate two versions of each of these two counterfactuals. In the first version, we assume that aggregate nation-wide labour supply is infinitely elastic. Hence, the population will adjust to the exogenous benchmark level of utility. All

[11]See for ranking public transport systems of cities:
https://www.businessinsider.com/cities-with-best-public-transportation-systems-2014-1?international=true&r=US&IR=T
https://smartasset.com/mortgage/best-cities-for-public-transportation

Table 4: Aggregate results for Counterfactuals

Counterfactual	No city form		No agglomeration			1979-2015		
Elasticity Labour Supply	∞	0	∞	0	(Level 79-15)	Model	Actual	(Level 1979)
Δlandrent (%)	-34.2	-12.5	-97.1	-56.0	2.03E+10	1499.9	540.1	5.3E+09
of which: cities ($\Delta, \%$)	-58.8	-45.2	-98.6	-78.5	1.18E+10	2567.1	643.0	2.52E+0.9
rural ($\Delta, \%$)	0	33.0	-95.1	-24.6	8.50E+09	546.2	449.3	2.82E+0.9
Δpopulation (%)	-12.4	0	-79.4	0	1.64E+08	–	44.8	1.42E+08
of which: cities ($\Delta, \%$)	-32.8	-21.5	-86.0	-30.1	6.21E+07	163.5	47.8	5.78E+07
rural ($\Delta, \%$)	0	16.8	-75.4	22.8	1.02E+08	46.35	42.3	8.44E+07
Δutility (income equi. $\Delta, \%$)	0	-2.9	0	-33.9	0.16	35.5	116.3	0.39
ΔH (market prices, %)	0	0	0	0	0.16	–	0.40	0.39

welfare effects fall upon the class of absentee landlords. In this version, ruling out the city form will have no effect on rural regions. In the second version, we assume that total labour supply is fixed but that the distribution of human capital adjusts to the shifts in demand. This implies that workers utility $u(h)$ will become endogenous. A simple way to deal with this counterfactual while still using the expressions in equation (19) is to change the Januari temperature T_r in all regions by the same amount. An increase in temperature is equivalent to a fall in the outside utility (because the increase in temperature is counterfactual: a worker would still get receive the same utility if the temperature would have increased, but is hasn't, making him worse of). We adjust the temperature till the total population in the counterfactual is the same as in reference point.

The aggregate results for this exercise are presented in Table 4. Results per region are presented in the Appendix.

Finally, we solve the model with the data for H_r in 1979 and 2015. We calculate a new reference point for 1979 based on the population in rural areas in that year, the data for H_r, and the model prediction for the population of cities in 1979. Next, we calculate a counterfactual using the actual growth of the total population from 1979 till 2015 and the data for H_r for 2015. The radius S_r of cities may shrink or grow compared from the reference point in 1979 to the endpoint in 2015 in this counterfactual. To avoid inconsistencies in the land endowment, we subtract the additional use of a city from the land endowment of the rural region of the state to which the city belongs, and the reverse when a city uses less land.[12]

5 Conclusion

TO BE COMPLETED

[12]For SMSA that belong to multiple states, we the state with the largest weight. We merge New Jersey to New York and DC to Maryland respectively since New Jersey and DC have no free rural available.

Appendix Derivation of ω_r^c and n_r^c

The population of a city satisfies

$$
\begin{aligned}
n_r^c &= \ln\left[\int_0^{S_r} 2\pi s e^{-l_r}\mathrm{d}s\right] = \ln\pi - 2\ln\kappa + 2\ln\left(\kappa S_r\right) - l_r \\
&= \lambda_\omega \omega_r^c - H_r + \lambda_\eta \lambda_z \alpha_T T_r + \ln\pi - 2\ln\kappa + 2\ln\left(\kappa S_r\right) \\
&= \Psi^{-1}\omega_r^r + \lambda_\omega \omega_r^c - \theta H_r + 2\Delta + 2\ln\left(\kappa S_r\right) + \ln 3.
\end{aligned}
$$

where substitute equation (7) for l_r and equation (16) for $\ln\delta - \ln\kappa$ in the second line. Using this result and equation (15) yields

$$
\begin{aligned}
\omega_r^c &= \psi\left(n_r^c + f_r + \theta H_r\right) \\
&= \psi\left[\Psi^{-1}\omega_r^r + \lambda_\omega \omega_r^c + 2\Delta + 2\ln\left(\kappa S_r\right) + \ln\left(3 - 2\kappa S_r\right)\right] \\
&= \omega_r^r + \Psi\left[2\Delta + 2\ln\left(1 - e^{-\omega_r^c}\right) + \ln\left(1 + 2e^{-\omega_r^c}\right)\right],
\end{aligned} \tag{29}
$$

where we substitute equation (12) for ω_r^r and equation (14) for κS_r. Hence

$$
\omega_r^r + 2\Psi\Delta = \omega_r^c - \Psi\left[2\ln\left(1 - e^{-\omega_r^c}\right) + \ln\left(1 + 2e^{-\omega_r^c}\right)\right]. \tag{30}
$$

Define the implicit function $\Gamma\left(\cdot\right)$ such that

$$
\begin{aligned}
\Gamma^{-1}(\omega_r^c) &\equiv \omega_r^c - \Psi\left[2\ln\left(1 - e^{-\omega_r^c}\right) + \ln\left(1 + 2e^{-\omega_r^c}\right)\right] \Rightarrow \\
\omega_r^c &= \Gamma\left(\omega_r^r + 2\Psi\Delta\right).
\end{aligned}
$$

Define the implicit function $\Gamma\left(\cdot\right)$ such that

$$
\begin{aligned}
\Gamma^{-1}(\omega_r^c) &\equiv \omega_r^c - \Psi\left[2\ln\left(1 - e^{-\omega_r^c}\right) + \ln\left(1 + 2e^{-\omega_r^c}\right)\right] \Rightarrow \\
\omega_r^c &= \Gamma\left[\omega_r^r + 2\Psi\Delta\right].
\end{aligned}
$$

The first line of equation (28) can be solved for $e^{-\overline{\omega}^c_{r\in\mathrm{C}}}$

$$
\begin{aligned}
\gamma \equiv \Gamma'\left[\overline{\omega}^r_{r\in\mathrm{C}} + 2\Psi\Delta\right] &= \frac{\mathrm{d}\overline{\omega}^c_{r\in\mathrm{C}}}{\mathrm{d}\overline{\omega}^r_{r\in\mathrm{C}}} \Rightarrow \\
\gamma^{-1} &= 1 - \frac{6\Psi e^{-2\overline{\omega}^c_{r\in\mathrm{C}}}}{\left(1 - e^{-\overline{\omega}^c_{r\in\mathrm{C}}}\right)\left(1 + 2e^{-\overline{\omega}^c_{r\in\mathrm{C}}}\right)} < 1 \Rightarrow \\
\Omega &= \frac{e^{-2\overline{\omega}^c_{r\in\mathrm{C}}}}{\left(1 - e^{-\overline{\omega}^c_{r\in\mathrm{C}}}\right)\left(1 + 2e^{-\overline{\omega}^c_{r\in\mathrm{C}}}\right)}, \qquad \Omega \equiv -\frac{\overline{\gamma}}{6\Psi\gamma} \Rightarrow \\
\overline{\omega}^c_{r\in\mathrm{C}} &= -\ln\left(\frac{\Omega + \sqrt{\Omega\left(9\Omega + 4\right)}}{4\Omega + 2}\right).
\end{aligned}
$$

Then, the second line can be solved for Δ

$$
\Delta \equiv \frac{\Delta_\Gamma}{2\Psi} - \left[\ln\left(1 - e^{-\overline{\omega}^c_{r\in\mathrm{C}}}\right) + \frac{1}{2}\ln\left(1 + 2e^{-\overline{\omega}^c_{r\in\mathrm{C}}}\right)\right].
$$

By equation (12), H_r can be written as

$$
H_r = -\left(\Psi\overline{\theta}\right)^{-1}\omega^r_r,
$$

where we omit terms that do not depend on H_r. Substitution in equation (18) yields

$$
\begin{aligned}
n^c_r &= \psi^{-1}\omega^c_r + \frac{\theta}{\Psi\overline{\theta}}\omega^r_r - \ln\left(1 + 2e^{-\omega^c_r}\right) + \ln 3 \\
&= \overline{\theta}^{-1}\left[\left(\psi^{-1} - \lambda_\omega\theta\right)\omega^c_r - \ln\left(1 + 2e^{-\omega^c_r}\right) - 2\theta\ln\left(1 - e^{-\omega^c_r}\right)\right] + cons.
\end{aligned}
$$

where we substitute ω^r_r for equation (30). The first two partial derivatives of n^c_r with respect to ω^c_r read

$$
\begin{aligned}
n^c_\omega &= \overline{\theta}^{-1}\left(\psi^{-1} - \lambda_\omega\theta + \frac{2e^{-\omega^c_r}}{1 + 2e^{-\omega^c_r}} - \frac{2\theta e^{-\omega^c_r}}{1 - e^{-\omega^c_r}}\right), \\
n^c_{\omega\omega} &= \overline{\theta}^{-1}\left(\frac{-2e^{-\omega^c_r}}{\left(1 + 2e^{-\omega^c_r}\right)^2} + \frac{2\theta e^{-\omega^c_r}}{\left(1 - e^{-\omega^c_r}\right)^2}\right),
\end{aligned}
$$

Using $\omega^c_H = -\gamma\Psi\overline{\theta}$, the partial derivatives of n^c_r with respect to H_r read

$$
\begin{aligned}
n^c_H &= -\gamma\Psi\overline{\theta}n^c_\omega = -\gamma\Psi\left(\psi^{-1} - \lambda_\omega\theta + \frac{2e^{-\omega^c_r}}{1 + 2e^{-\omega^c_r}} - \frac{2\theta e^{-\omega^c_r}}{1 - e^{-\omega^c_r}}\right) = 7.41, \\
n^c_{HH} &= \left(\gamma\Psi\overline{\theta}\right)^2 n^c_{\omega\omega} = -\left(\gamma\Psi\right)^2\left(\theta - 1\right)\left(\frac{-2e^{-\omega^c_r}}{\left(1 + 2e^{-\omega^c_r}\right)^2} + \frac{2\theta e^{-\omega^c_r}}{\left(1 - e^{-\omega^c_r}\right)^2}\right) = -18.23, \\
\Omega &= 0.72, \quad \omega^c_r = 0.34.
\end{aligned} \tag{31}
$$

where the numerical values follows $\Psi = 0.12, \psi = 0.08, \lambda_\omega = 3.9, \theta = 5$, and $\gamma = 2$.

Table A1: CBSA Observations Distribution Among States

CBSA	State I	State II	State III	State IV	Pct SI	Pct SII	Pct SIII	Pct SIV	NAME
31100	CA				100.00%				Los Angeles-Long Beach-Anaheim, CA
40140	CA				100.00%				Riverside-San Bernardino-Ontario, CA
41740	CA				100.00%				San Diego-Carlsbad, CA
41860	CA				100.00%				San Francisco-Oakland-Hayward, CA
41940	CA				100.00%				San Jose-Sunnyvale-Santa Clara, CA
19740	CO				100.00%				Denver-Aurora-Lakewood, CO
47900	DC	VA	MD		45.91%	25.90%	28.19%		Washington-Arlington-Alexandria, DC-VA-MD-WV
33100	FL				100.00%				Miami-Fort Lauderdale-West Palm Beach, FL
45300	FL				100.00%				Tampa-St. Petersburg-Clearwater, FL
12060	GA				100.00%				Atlanta-Sandy Springs-Roswell, GA
16980	IL	IN	WI		98.23%	1.77%	0.00%		Chicago-Naperville-Elgin, IL-IN-WI
26900	IN				100.00%				Indianapolis-Carmel-Anderson, IN
35380	LA				100.00%				New Orleans-Metairie, LA
14460	MA	NH			86.75%	13.25%			Boston-Cambridge-Newton, MA-NH
12580	MD				100.00%				Baltimore-Columbia-Towson, MD
19820	MI				100.00%				Detroit-Warren-Dearborn, MI
33460	MN	WI			99.99%	0.01%			Minneapolis-St. Paul-Bloomington, MN-WI
28140	MO	KS			45.36%	54.64%			Kansas City, MO-KS
41180	MO	IL			80.98%	19.02%			St. Louis, MO-IL
24660	NC				100.00%				Greensboro-High Point, NC
15380	NY				100.00%				Buffalo-Cheektowaga-Niagara Falls, NY
35620	NY	NJ			69.24%	30.76%			New York-Newark-Jersey City, NY-NJ
40380	NY				100.00%				Rochester, NY
17140	OH	KY			77.70%				Cincinnati, OH-KY-IN
17460	OH				100.00%				Cleveland-Elyria, OH
18140	OH				100.00%				Columbus, OH
38900	OR	WA			91.57%	8.43%			Portland-Vancouver-Hillsboro, OR-WA
37980	PA	NJ	DE	MD	62.06%	23.32%	14.62%	0.00%	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
38300	PA				100.00%				Pittsburgh, PA
19100	TX				100.00%				Dallas-Fort Worth-Arlington, TX
26420	TX				100.00%				Houston-The Woodlands-Sugar Land, TX
47260	VA				100.00%				Virginia Beach-Norfolk-Newport News, VA-NC
42660	WA				100.00%				Seattle-Tacoma-Bellevue, WA
33340	WI				100.00%				Milwaukee-Waukesha-West Allis, WI

Note: Information for 34 city areas: CBSA code in 2013, city belong to which state(s) and the percentage of sample observations in the CPS 1979-2015, name of cities. *Data sources*: the Current Population Survey MORG and the US Census Bureau.

Table A2: Individual Mincerian Wage Regression

Variables	Coefficient	t-stat	Variables	Coefficient	t-stat
Male	0.419	(448.59)	Edu = 0	-0.652	(-98.12)
Male $\times$ Time Trend	-0.00565	(-118.75)	Edu = 1	-0.532	(-36.40)
Single	0.0516	(40.86)	Edu = 2	-0.540	(-76.20)
Male $\times$ Single	-0.283	(-172.09)	Edu = 3	-0.532	(-86.27)
Single $\times$ Time Trend	-0.00287	(-48.72)	Edu = 4	-0.452	(-73.50)
Male $\times$ Single $\times$ Time Trend	0.00449	(55.31)	Edu = 5	-0.475	(-113.51)
Divorced	0.0213	(14.06)	Edu = 6	-0.426	(-129.18)
Male $\times$ Divorced	-0.107	(-41.03)	Edu = 7	-0.357	(-115.49)
Divorced $\times$ Time Trend	-0.00175	(-0.196)	Edu = 8	-0.264	(-128.79)
Male $\times$ Divorced $\times$ Time Trend	0.00185	(15.19)	Edu = 9	-0.260	(-180.71)
South	0.00858	(2.58)	Edu = 10	-0.194	(-188.11)
Black	-0.0975	(-96.7)	Edu = 11	-0.157	(-167.93)
Black $\times$ South	-0.0382	(-28.34)	Edu = 13	0.0611	(67.86)
Other Race	-0.0834	(-69.98)	Edu = 14	0.168	(193.93)
Other $\times$ South	-0.00692	(-2.75)	Edu = 15	0.212	(136.26)
Year of Experience	0.0299	(51.92)	Edu = 16	0.420	(373.93)
Exp $\times$ Edu	0.00151	(33.78)	Edu = 17	0.405	(171.13)
Exp^2 / 100	-0.0445	(-16.72)	Edu = 18	0.576	(326.16)
Exp^2 / 100 $\times$ Edu	-0.00842	(-40.15)	Constant	1.109	(267.11)
Exp^3 / 100000	0.183	(5.25)			
Exp^3 / 100000 $\times$ Edu	0.102	(36.01)	Observations	5,426,947	
Edu in y9297	0.00463	(21.76)	R-squared	0.575	
			R-MSE	0.444	

Note: Table presents the estimated β using OLS regression. Dependent variable is the log hourly wage. Mincer wage regression includes individual characteristics x, gender, year of education, year of experience, race, marital status, and the interaction of these factors. All the regressions include time x region dummies. Robust t-statistics in parentheses.

Table A3: Ranking of Regions in Human Capital Level

Region	HC Index	Type	Region	HC Index	Type
Boston, MA	0.139	City	North Dakota	0.005	Non-city
San Jose, CA	0.128	City	Kansas	0.003	Non-city
San Francisco, CA	0.128	City	Nebraska	-0.002	Non-city
Seattle, WA	0.122	City	Oklahoma	-0.002	Non-city
Connecticut	0.104	Non-city	Illinois	-0.004	Non-city
Portland, OR	0.103	City	Iowa	-0.005	Non-city
Denver, CO	0.101	City	Pennsylvania	-0.008	Non-city
Washington, DC	0.093	City	Arizona	-0.008	Non-city
Pittsburgh, PA	0.088	City	Utah	-0.009	Non-city
Minneapolis, MN	0.084	City	Idaho	-0.010	Non-city
Rochester, NY	0.073	City	Wisconsin	-0.011	Non-city
New Hampshire	0.064	Non-city	Dallas, TX	-0.012	City
Colorado	0.063	Non-city	Ohio	-0.013	Non-city
New York, NY	0.062	City	Kentucky	-0.013	Non-city
Kansas City, MO	0.061	City	Florida	-0.013	Non-city
Vermont	0.057	Non-city	Delaware	-0.015	Non-city
Philadelphia, PA	0.054	City	Miami, FL	-0.019	City
New York	0.047	Non-city	Houston, TX	-0.025	City
Milwaukee, WI	0.046	City	Greensboro, NC	-0.026	City
Chicago, IL	0.045	City	Minnesota	-0.026	Non-city
Indianapolis, IN	0.044	City	New Orleans, LA	-0.027	City
San Diego, CA	0.042	City	California	-0.030	Non-city
Baltimore, MD	0.037	City	South Dakota	-0.031	Non-city
Massachusetts	0.036	Non-city	Indiana	-0.031	Non-city
Detroit, MI	0.034	City	Virginia Beach, VA	-0.031	City
Cleveland, OH	0.033	City	Virginia	-0.036	Non-city
Atlanta, GA	0.033	City	Missouri	-0.037	Non-city
Montana	0.031	Non-city	North Carolina	-0.038	Non-city
St Louis, MO	0.027	City	Nevada	-0.040	Non-city
Maine	0.027	Non-city	Tennessee	-0.040	Non-city
Buffalo, NY	0.026	City	Alabama	-0.044	Non-city
West Virginia	0.025	Non-city	Los Angeles, CA	-0.044	City
Washington	0.022	Non-city	Riverside, CA	-0.047	City
Columbus, OH	0.021	City	Maryland	-0.053	Non-city
Wyoming	0.020	Non-city	South Carolina	-0.060	Non-city
Cincinnati, OH	0.016	City	Texas	-0.067	Non-city
New Mexico	0.015	Non-city	Arkansas	-0.076	Non-city
Rhode Island	0.011	Non-city	Louisiana	-0.080	Non-city
Michigan	0.008	Non-city	Mississippi	-0.081	Non-city
Oregon	0.008	Non-city	Georgia	-0.122	Non-city
Tampa, FL	0.007	City			

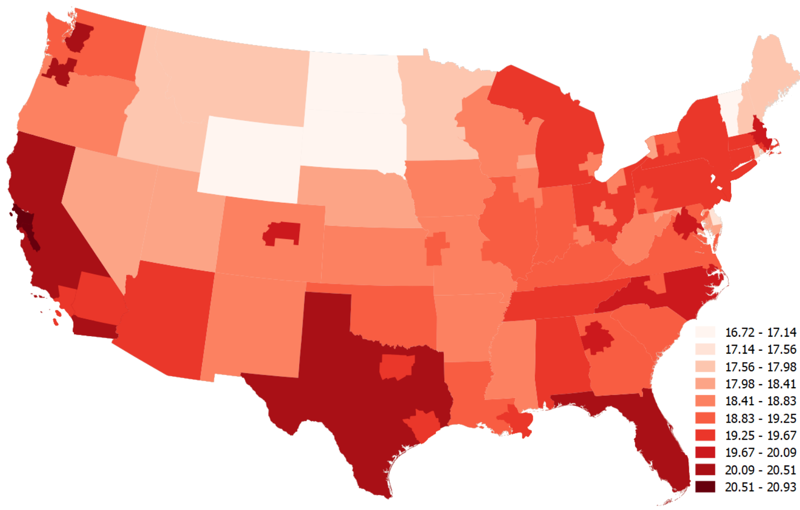
Note: Average local human capital index and region type. 34 cities are denoted by the name of largest city with the abbreviation of the state. 47 non-city areas are denoted by the name of the states. Detailed definitions of occupation index in section 2. *Data sources*: Current Population Survey MORG and author's own calculations.

Table A4: Counterfactual Results with Perfect Elastic Labour Supply - I

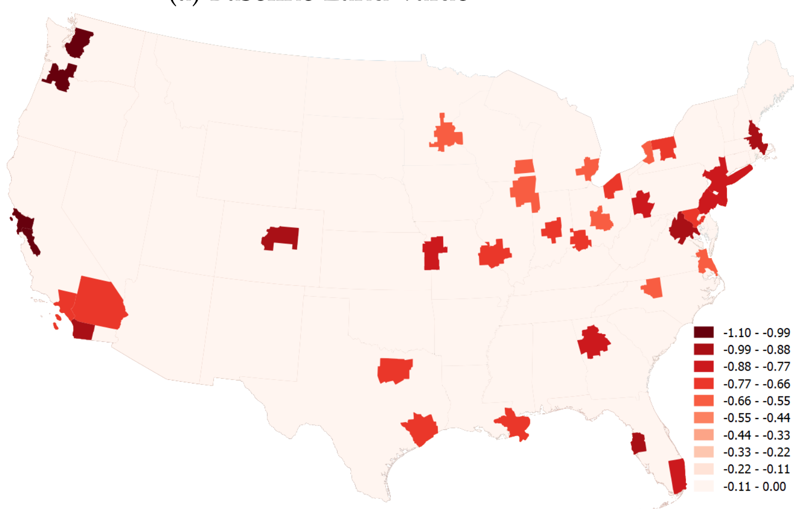
Region	Log Population	No City	No Agg	log Land Value	No City	No Agg
Atlanta, GA	14.55	-0.40	-1.97	19.75	-0.85	-4.19
Baltimore, MD	14.20	-0.36	-1.87	19.19	-0.76	-3.98
Boston, MA	14.57	-0.46	-2.13	19.91	-0.98	-4.54
Buffalo, NY	13.68	-0.28	-1.69	18.36	-0.59	-3.61
Chicago, IL	13.71	-0.29	-1.73	18.42	-0.63	-3.69
Cincinnati, OH	13.98	-0.31	-1.77	18.81	-0.67	-3.77
Cleveland, OH	13.97	-0.32	-1.79	18.82	-0.69	-3.82
Columbus, OH	13.90	-0.30	-1.75	18.69	-0.65	-3.73
Dallas, TX	14.42	-0.35	-1.86	19.48	-0.75	-3.96
Denver, CO	14.47	-0.43	-2.04	19.70	-0.91	-4.35
Detroit, MI	13.75	-0.29	-1.73	18.47	-0.62	-3.68
Greensboro, NC	14.03	-0.29	-1.72	18.86	-0.62	-3.67
Houston, TX	14.52	-0.36	-1.87	19.63	-0.77	-3.98
Indianapolis, IN	14.00	-0.34	-1.82	18.88	-0.72	-3.87
Kansas City, MO	14.20	-0.37	-1.90	19.22	-0.79	-4.05
Los Angeles, CA	14.50	-0.34	-1.84	19.58	-0.73	-3.91
Miami, FL	14.89	-0.41	-1.99	20.22	-0.87	-4.24
Milwaukee, WI	13.65	-0.29	-1.72	18.33	-0.61	-3.65
Minneapolis, MN	13.32	-0.26	-1.67	17.85	-0.56	-3.56
New Orleans, LA	14.52	-0.36	-1.87	19.62	-0.76	-3.97
New York, NY	14.34	-0.39	-1.95	19.44	-0.83	-4.14
Philadelphia, PA	14.28	-0.38	-1.92	19.33	-0.81	-4.08
Pittsburgh, PA	14.06	-0.37	-1.90	19.03	-0.79	-4.05
Portland, OR	14.82	-0.47	-2.15	20.27	-1.00	-4.58
Riverside, CA	14.43	-0.33	-1.81	19.46	-0.71	-3.86
Rochester, NY	14.00	-0.36	-1.86	18.91	-0.76	-3.96
St Louis, MO	14.07	-0.33	-1.81	18.97	-0.71	-3.86
San Diego, CA	14.95	-0.45	-2.10	20.39	-0.96	-4.47
San Francisco, CA	15.20	-0.52	-2.30	20.92	-1.10	-4.90
San Jose, CA	15.20	-0.52	-2.30	20.93	-1.10	-4.90
Seattle, WA	14.95	-0.49	-2.22	20.51	-1.04	-4.72
Tampa, FL	14.89	-0.42	-2.03	20.26	-0.90	-4.32
Virginia Beach, VA	14.06	-0.29	-1.72	18.89	-0.62	-3.66
Washington, DC	14.58	-0.44	-2.07	19.87	-0.93	-4.40
Maine	13.62	0.00	-1.30	17.67	0.00	-2.76
New Hampshire	13.57	0.00	-1.42	17.87	0.00	-3.02
Vermont	12.88	0.00	-1.37	17.07	0.00	-2.92
Massachusetts	14.90	0.00	-1.44	19.31	0.00	-3.07
Rhode Island	13.43	0.00	-1.43	17.85	0.00	-3.04
Connecticut	14.63	0.00	-1.58	19.27	0.00	-3.36

Table A5: Counterfactual Results with Perfect Elastic Labour Supply - II

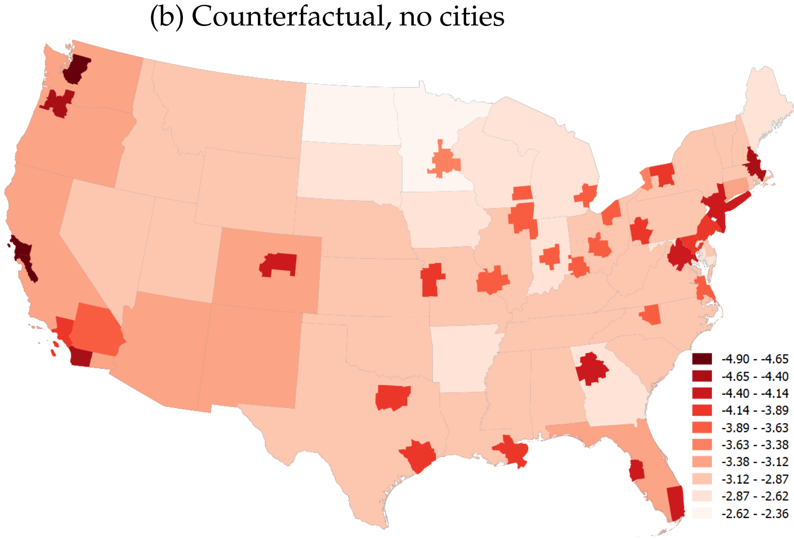
Region	Log Population	No City	No Agg	log Land Value	No City	No Agg
New York	14.95	0.00	-1.44	19.32	0.00	-3.06
Pennsylvania	15.16	0.00	-1.37	19.47	0.00	-2.92
Ohio	15.18	0.00	-1.37	19.48	0.00	-2.92
Indiana	14.88	0.00	-1.33	19.12	0.00	-2.84
Illinois	14.79	0.00	-1.37	19.09	0.00	-2.92
Michigan	15.43	0.00	-1.34	19.62	0.00	-2.85
Wisconsin	14.68	0.00	-1.24	18.64	0.00	-2.63
Minnesota	14.06	0.00	-1.11	17.71	0.00	-2.36
Iowa	14.44	0.00	-1.30	18.55	0.00	-2.76
Missouri	14.24	0.00	-1.35	18.53	0.00	-2.87
North Dakota	12.97	0.00	-1.17	16.72	0.00	-2.49
South Dakota	13.06	0.00	-1.25	17.08	0.00	-2.66
Nebraska	13.90	0.00	-1.36	18.17	0.00	-2.90
Kansas	14.02	0.00	-1.42	18.44	0.00	-3.03
Delaware	13.12	0.00	-1.42	17.56	0.00	-3.02
Maryland	13.92	0.00	-1.34	18.23	0.00	-2.86
Virginia	14.61	0.00	-1.39	19.01	0.00	-2.95
West Virginia	13.99	0.00	-1.46	18.47	0.00	-3.12
North Carolina	15.37	0.00	-1.41	19.83	0.00	-3.01
South Carolina	14.76	0.00	-1.40	19.23	0.00	-2.98
Georgia	14.73	0.00	-1.29	19.02	0.00	-2.74
Florida	15.54	0.00	-1.56	20.35	0.00	-3.33
Kentucky	14.69	0.00	-1.42	19.13	0.00	-3.03
Tennessee	15.11	0.00	-1.39	19.52	0.00	-2.96
Alabama	14.86	0.00	-1.43	19.38	0.00	-3.04
Mississippi	14.39	0.00	-1.36	18.80	0.00	-2.90
Arkansas	14.32	0.00	-1.34	18.66	0.00	-2.85
Louisiana	14.53	0.00	-1.39	19.01	0.00	-2.96
Oklahoma	14.61	0.00	-1.47	19.15	0.00	-3.12
Texas	15.73	0.00	-1.40	20.22	0.00	-2.98
Montana	13.27	0.00	-1.40	17.57	0.00	-2.98
Idaho	13.57	0.00	-1.35	17.83	0.00	-2.88
Wyoming	12.72	0.00	-1.37	16.98	0.00	-2.92
Colorado	14.11	0.00	-1.49	18.60	0.00	-3.17
New Mexico	13.92	0.00	-1.48	18.48	0.00	-3.16
Arizona	14.89	0.00	-1.50	19.52	0.00	-3.19
Utah	14.09	0.00	-1.37	18.39	0.00	-2.92
Nevada	13.92	0.00	-1.36	18.25	0.00	-2.89
Washington	14.49	0.00	-1.50	19.07	0.00	-3.18
Oregon	13.94	0.00	-1.48	18.49	0.00	-3.14
California	15.58	0.00	-1.47	20.19	0.00	-3.13



(a) Baseline Land Value

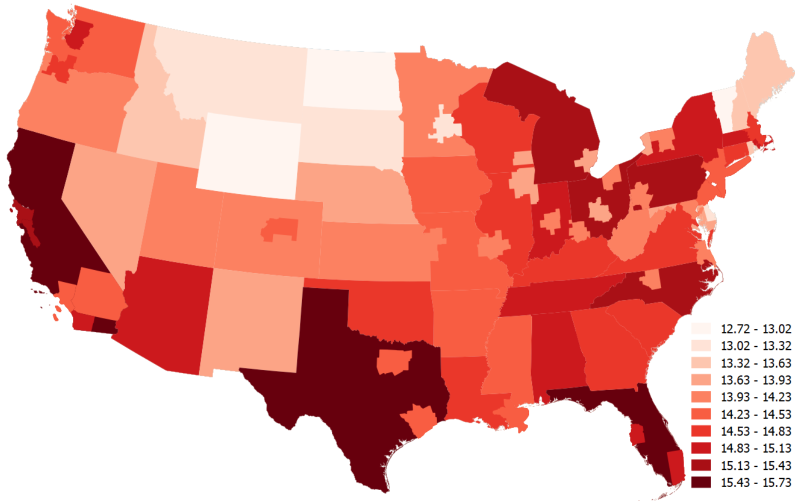


(b) Counterfactual, no cities

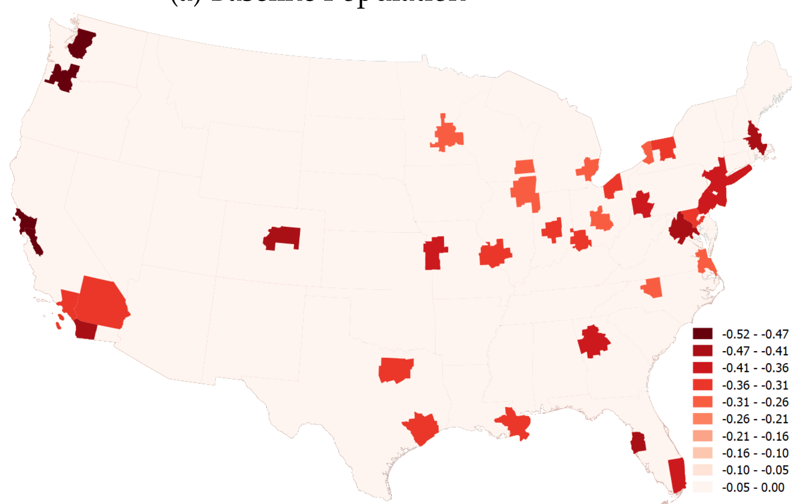


(c) Counterfactual, no agglomeration

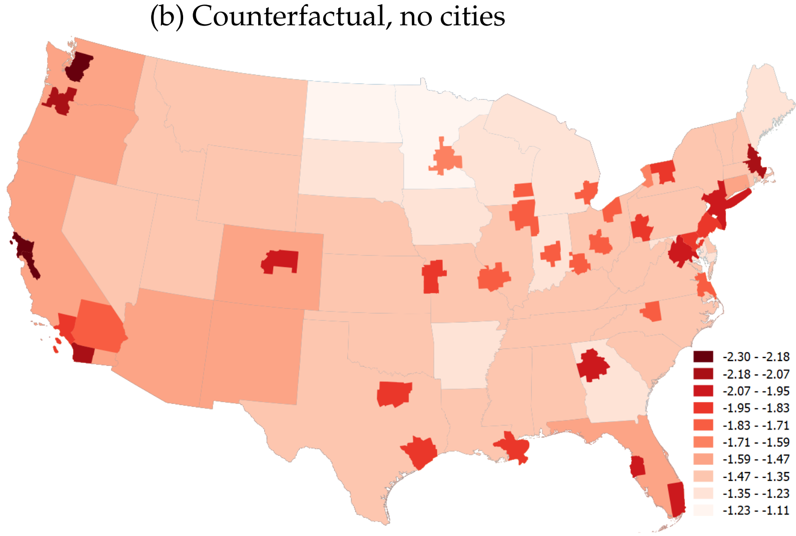
Figure A1: Average Land Value and Counterfactuals with Elastic Labour Supply



(a) Baseline Population



(b) Counterfactual, no cities



(c) Counterfactual, no agglomeration

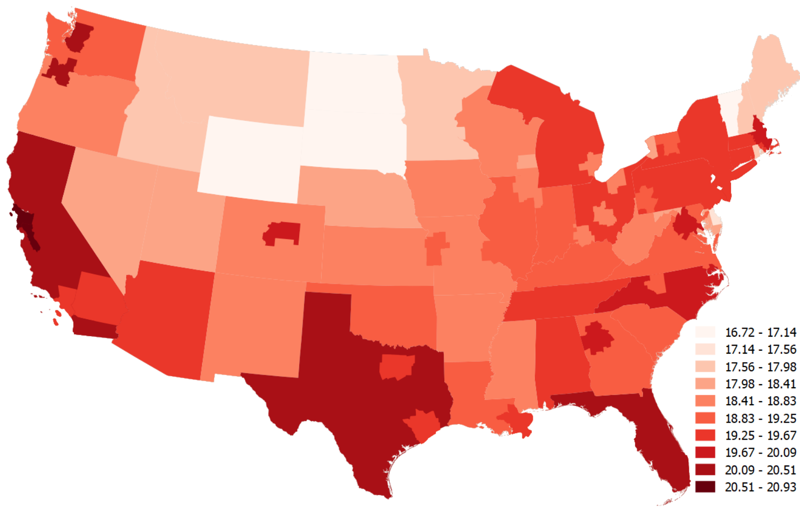
Figure A2: Average Population and Counterfactuals with Elastic Labour Supply

Table A6: Counterfactual Results with Perfect Inelastic Labour Supply - I

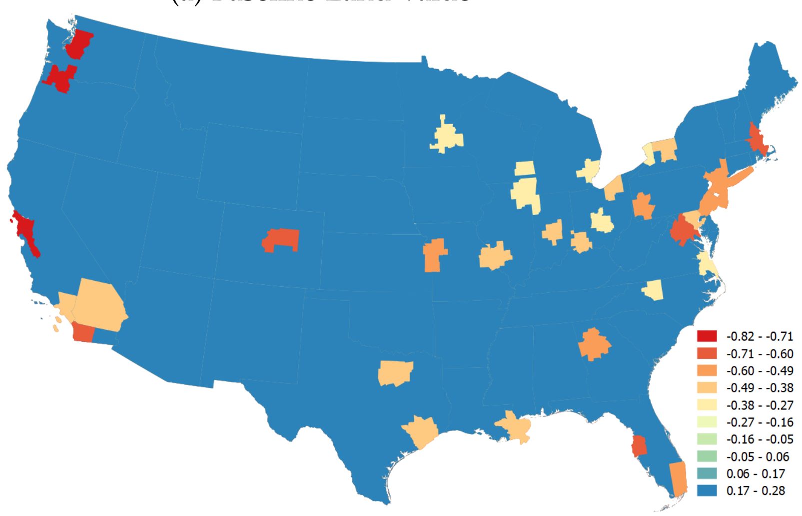
Region	Log Population	No City	No Agg	log Land Value	No City	No Agg
Atlanta, GA	14.55	-0.24	-0.36	19.75	-0.57	-1.46
Baltimore, MD	14.20	-0.20	-0.26	19.19	-0.48	-1.25
Boston, MA	14.57	-0.31	-0.52	19.91	-0.70	-1.81
Buffalo, NY	13.68	-0.12	-0.09	18.36	-0.30	-0.88
Chicago, IL	13.71	-0.14	-0.13	18.42	-0.34	-0.96
Cincinnati, OH	13.98	-0.16	-0.16	18.81	-0.38	-1.04
Cleveland, OH	13.97	-0.17	-0.19	18.82	-0.41	-1.09
Columbus, OH	13.90	-0.15	-0.15	18.69	-0.36	-1.01
Dallas, TX	14.42	-0.20	-0.25	19.48	-0.47	-1.24
Denver, CO	14.47	-0.27	-0.44	19.70	-0.63	-1.63
Detroit, MI	13.75	-0.14	-0.12	18.47	-0.34	-0.95
Greensboro, NC	14.03	-0.13	-0.12	18.86	-0.33	-0.95
Houston, TX	14.52	-0.20	-0.26	19.63	-0.48	-1.26
Indianapolis, IN	14.00	-0.18	-0.21	18.88	-0.43	-1.15
Kansas City, MO	14.20	-0.22	-0.30	19.22	-0.51	-1.33
Los Angeles, CA	14.50	-0.19	-0.23	19.58	-0.45	-1.19
Miami, FL	14.89	-0.25	-0.38	20.22	-0.59	-1.51
Milwaukee, WI	13.65	-0.13	-0.11	18.33	-0.33	-0.93
Minneapolis, MN	13.32	-0.11	-0.07	17.85	-0.28	-0.84
New Orleans, LA	14.52	-0.20	-0.26	19.62	-0.48	-1.25
New York, NY	14.34	-0.24	-0.34	19.44	-0.55	-1.42
Philadelphia, PA	14.28	-0.22	-0.31	19.33	-0.52	-1.36
Pittsburgh, PA	14.06	-0.22	-0.30	19.03	-0.51	-1.33
Portland, OR	14.82	-0.31	-0.55	20.27	-0.71	-1.86
Riverside, CA	14.43	-0.18	-0.21	19.46	-0.42	-1.13
Rochester, NY	14.00	-0.20	-0.26	18.91	-0.47	-1.24
St Louis, MO	14.07	-0.18	-0.21	18.97	-0.42	-1.14
San Diego, CA	14.95	-0.29	-0.49	20.39	-0.67	-1.74
San Francisco, CA	15.20	-0.36	-0.70	20.92	-0.82	-2.18
San Jose, CA	15.20	-0.36	-0.70	20.93	-0.82	-2.18
Seattle, WA	14.95	-0.34	-0.61	20.51	-0.76	-2.00
Tampa, FL	14.89	-0.27	-0.42	20.26	-0.62	-1.60
Virginia Beach, VA	14.06	-0.13	-0.12	18.89	-0.33	-0.94
Washington, DC	14.58	-0.28	-0.46	19.87	-0.65	-1.67
Maine	13.62	0.16	0.31	17.67	0.28	-0.04
New Hampshire	13.57	0.16	0.19	17.87	0.28	-0.30
Vermont	12.88	0.16	0.23	17.07	0.28	-0.20
Massachusetts	14.90	0.16	0.16	19.31	0.28	-0.35
Rhode Island	13.43	0.16	0.18	17.85	0.28	-0.32
Connecticut	14.63	0.16	0.03	19.27	0.28	-0.64

Table A7: Counterfactual Results with Perfect Inelastic Labour Supply - II

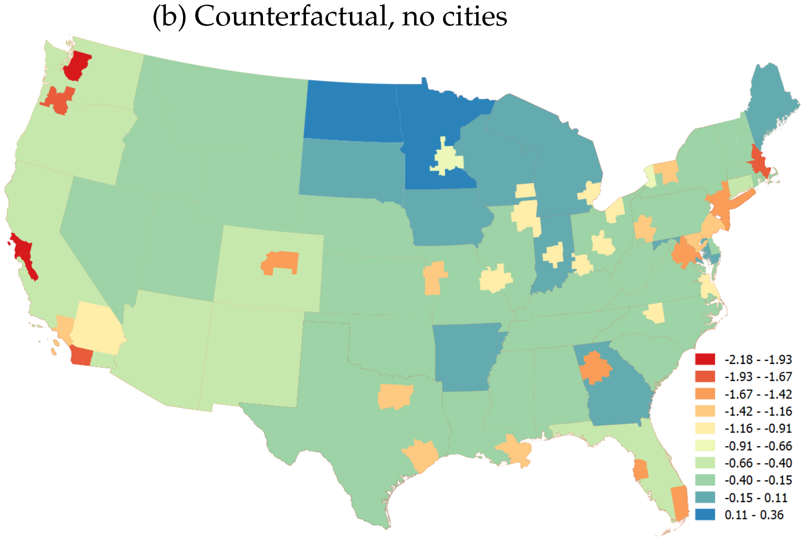
Region	Log Population	No City	No Agg	log Land Value	No City	No Agg
New York	14.95	0.16	0.17	19.32	0.28	-0.33
Pennsylvania	15.16	0.16	0.23	19.47	0.28	-0.20
Ohio	15.18	0.16	0.24	19.48	0.28	-0.19
Indiana	14.88	0.16	0.27	19.12	0.28	-0.11
Illinois	14.79	0.16	0.23	19.09	0.28	-0.20
Michigan	15.43	0.16	0.27	19.62	0.28	-0.13
Wisconsin	14.68	0.16	0.37	18.64	0.28	0.09
Minnesota	14.06	0.16	0.50	17.71	0.28	0.36
Iowa	14.44	0.16	0.31	18.55	0.28	-0.04
Missouri	14.24	0.16	0.26	18.53	0.28	-0.15
North Dakota	12.97	0.16	0.44	16.72	0.28	0.23
South Dakota	13.06	0.16	0.36	17.08	0.28	0.07
Nebraska	13.90	0.16	0.24	18.17	0.28	-0.18
Kansas	14.02	0.16	0.18	18.44	0.28	-0.31
Delaware	13.12	0.16	0.19	17.56	0.28	-0.30
Maryland	13.92	0.16	0.26	18.23	0.28	-0.14
Virginia	14.61	0.16	0.22	19.01	0.28	-0.23
West Virginia	13.99	0.16	0.14	18.47	0.28	-0.39
North Carolina	15.37	0.16	0.19	19.83	0.28	-0.29
South Carolina	14.76	0.16	0.21	19.23	0.28	-0.25
Georgia	14.73	0.16	0.32	19.02	0.28	-0.02
Florida	15.54	0.16	0.04	20.35	0.28	-0.61
Kentucky	14.69	0.16	0.18	19.13	0.28	-0.30
Tennessee	15.11	0.16	0.22	19.52	0.28	-0.24
Alabama	14.86	0.16	0.18	19.38	0.28	-0.32
Mississippi	14.39	0.16	0.24	18.80	0.28	-0.18
Arkansas	14.32	0.16	0.27	18.66	0.28	-0.13
Louisiana	14.53	0.16	0.22	19.01	0.28	-0.23
Oklahoma	14.61	0.16	0.14	19.15	0.28	-0.40
Texas	15.73	0.16	0.20	20.22	0.28	-0.26
Montana	13.27	0.16	0.21	17.57	0.28	-0.25
Idaho	13.57	0.16	0.25	17.83	0.28	-0.16
Wyoming	12.72	0.16	0.23	16.98	0.28	-0.20
Colorado	14.11	0.16	0.11	18.60	0.28	-0.45
New Mexico	13.92	0.16	0.12	18.48	0.28	-0.44
Arizona	14.89	0.16	0.11	19.52	0.28	-0.47
Utah	14.09	0.16	0.23	18.39	0.28	-0.20
Nevada	13.92	0.16	0.25	18.25	0.28	-0.17
Washington	14.49	0.16	0.11	19.07	0.28	-0.46
Oregon	13.94	0.16	0.13	18.49	0.28	-0.42
California	15.58	0.16	0.13	20.19	0.28	-0.41



(a) Baseline Land Value

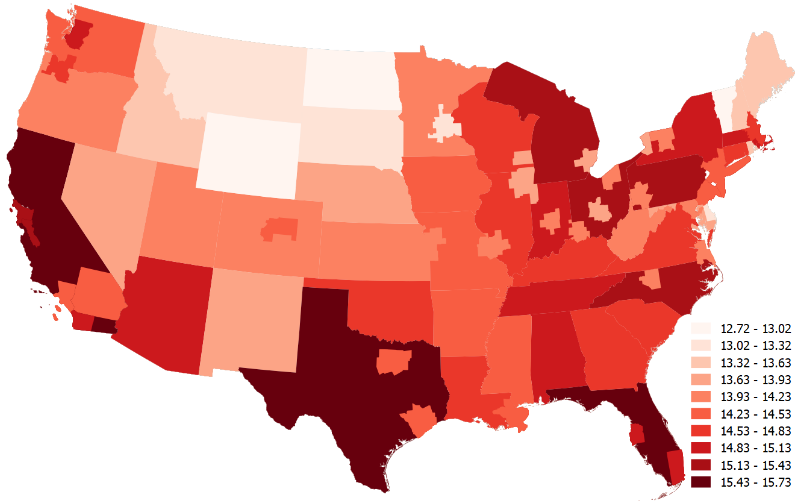


(b) Counterfactual, no cities

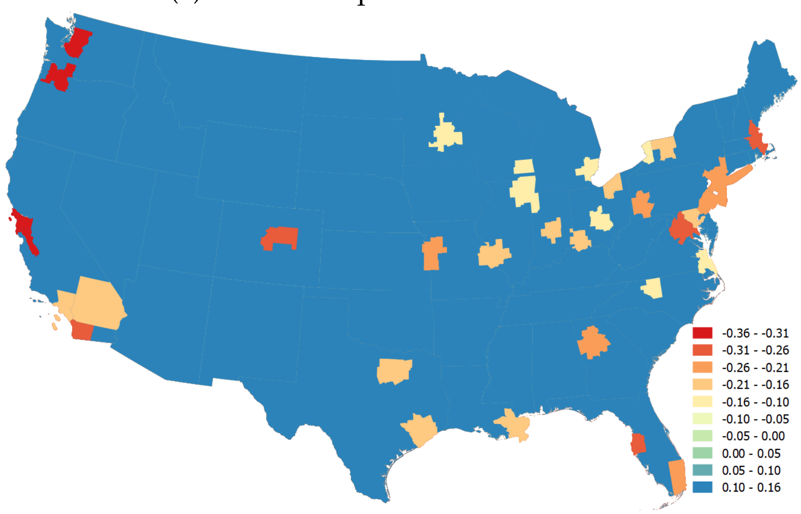


(c) Counterfactual, no agglomeration

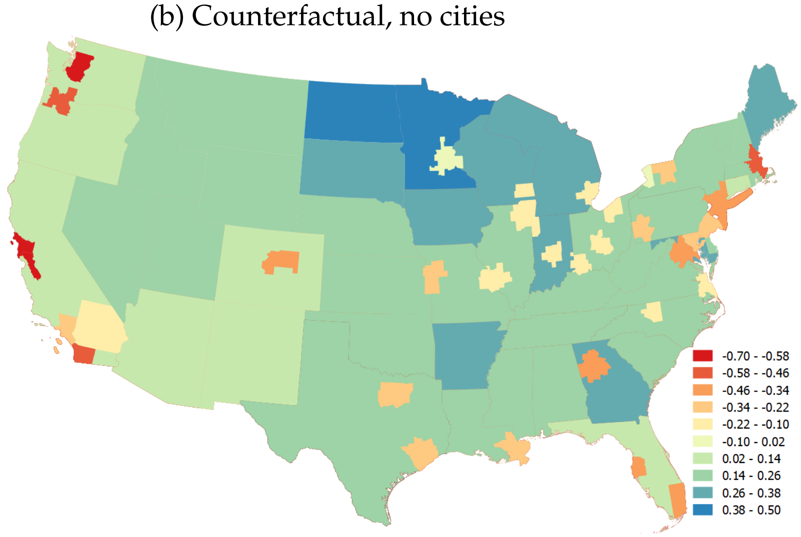
Figure A3: Average Land Value and Counterfactuals with Inelastic Labour Supply



(a) Baseline Population



(b) Counterfactual, no cities



(c) Counterfactual, no agglomeration

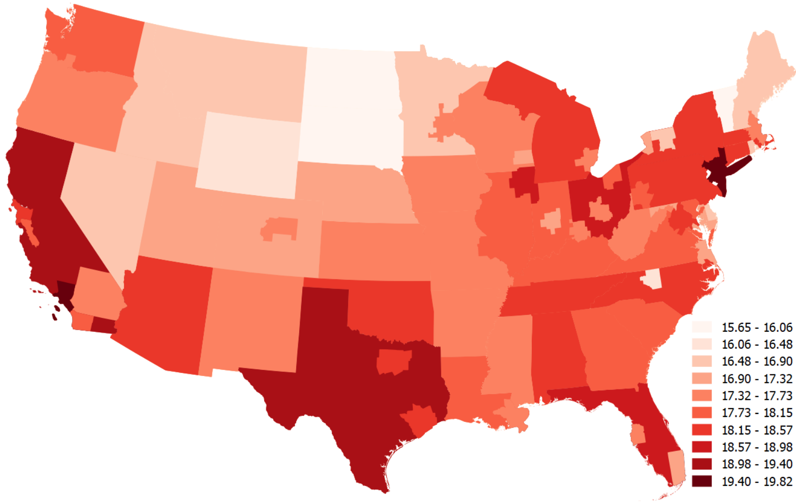
Figure A4: Average Population and Counterfactuals with Inelastic Labour Supply

Table A8: Counterfactual Results with Different Human Capital - I

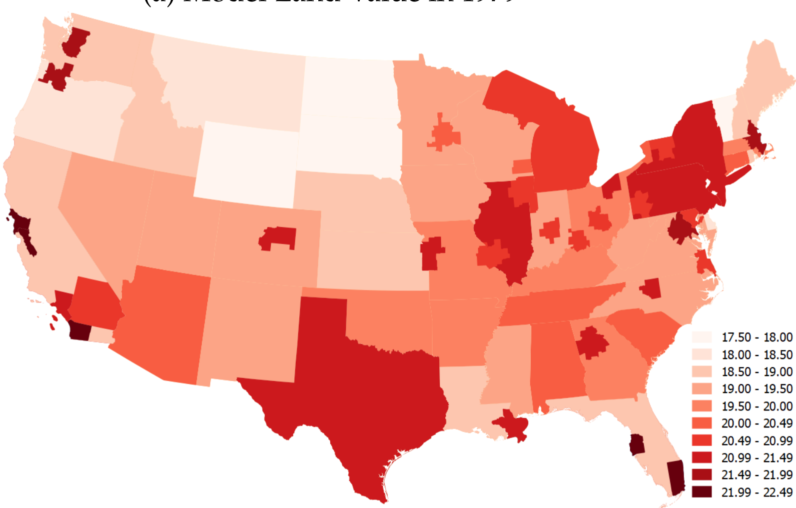
Region	Log Population	No City	No Agg	log Land Value	No City	No Agg
Atlanta, GA	14.25	15.37	1.12	18.10	21.37	3.26
Baltimore, MD	14.23	15.13	0.91	17.88	20.99	3.11
Boston, MA	13.93	15.38	1.45	17.55	21.59	4.04
Buffalo, NY	13.64	14.67	1.03	17.12	20.17	3.05
Chicago, IL	15.33	14.83	-0.49	18.87	20.53	1.66
Cincinnati, OH	13.95	15.04	1.09	17.41	20.82	3.41
Cleveland, OH	14.19	15.14	0.95	17.77	21.07	3.30
Columbus, OH	13.67	15.00	1.33	17.32	20.78	3.45
Dallas, TX	14.50	15.37	0.87	18.36	21.31	2.95
Denver, CO	13.81	15.16	1.35	17.60	21.07	3.47
Detroit, MI	14.25	14.83	0.58	17.67	20.50	2.83
Greensboro, NC	12.71	15.26	2.56	16.47	21.18	4.71
Houston, TX	14.54	15.40	0.85	18.36	21.31	2.95
Indianapolis, IN	13.59	15.03	1.44	17.00	20.85	3.85
Kansas City, MO	13.80	15.21	1.41	17.38	21.21	3.83
Los Angeles, CA	15.66	15.39	-0.27	19.66	21.27	1.60
Miami, FL	13.31	15.86	2.55	17.08	22.17	5.09
Milwaukee, WI	13.73	14.70	0.97	17.10	20.28	3.18
Minneapolis, MN	14.19	14.49	0.30	17.43	20.02	2.60
New Orleans, LA	13.63	15.38	1.74	17.67	21.27	3.59
New York, NY	16.10	15.28	-0.82	19.82	21.31	1.49
Philadelphia, PA	14.71	15.19	0.48	18.39	21.13	2.73
Pittsburgh, PA	14.39	14.98	0.59	18.07	20.82	2.75
Portland, OR	13.68	15.51	1.83	17.61	21.69	4.08
Riverside, CA	13.77	15.16	1.39	17.57	20.81	3.24
Rochester, NY	13.38	14.88	1.50	16.87	20.60	3.73
St Louis, MO	14.31	15.08	0.77	17.89	20.90	3.01
San Diego, CA	14.05	15.77	1.72	18.06	22.08	4.01
San Francisco, CA	13.65	15.93	2.28	18.19	22.49	4.31
San Jose, CA	13.72	15.88	2.16	17.90	22.38	4.48
Seattle, WA	13.89	15.63	1.74	18.03	21.93	3.90
Tampa, FL	13.77	15.86	2.09	17.62	22.23	4.61
Virginia Beach, VA	13.60	15.12	1.52	17.20	20.87	3.67
Washington, DC	14.46	15.60	1.14	18.55	21.97	3.42
Maine	13.49	13.68	0.18	16.57	18.53	1.96
New Hampshire	13.31	13.71	0.40	16.66	18.65	1.99
Vermont	12.71	12.95	0.23	15.94	17.93	1.99
Massachusetts	14.81	14.69	-0.12	18.23	19.91	1.68
Rhode Island	13.36	13.48	0.12	16.86	18.57	1.70
Connecticut	14.55	14.69	0.14	18.16	20.03	1.86

Table A9: Counterfactual Results with Different Human Capital - II

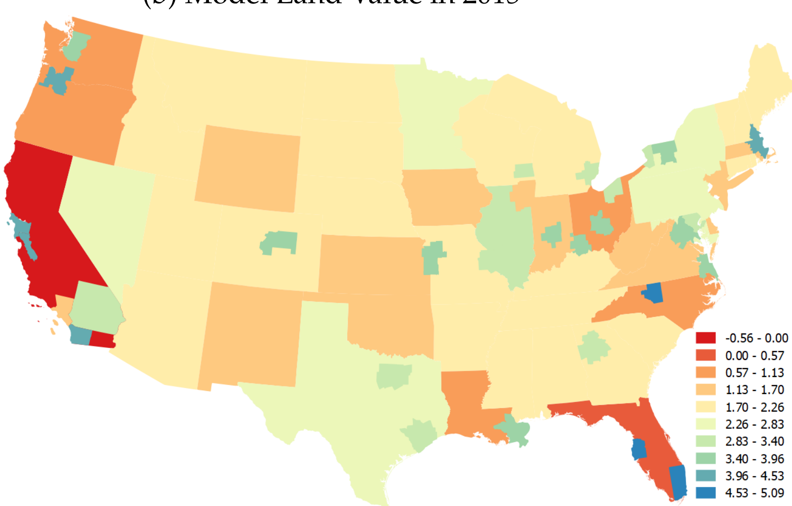
Region	Log Population	No City	No Agg	log Land Value	No City	No Agg
New York	14.89	16.23	1.34	18.51	21.21	2.69
Pennsylvania	15.10	16.17	1.07	18.56	21.06	2.50
Ohio	15.17	14.85	-0.33	18.66	19.68	1.02
Indiana	14.83	14.55	-0.27	18.13	19.39	1.26
Illinois	14.76	16.29	1.53	18.11	21.24	3.13
Michigan	15.33	15.91	0.58	18.55	20.76	2.21
Wisconsin	14.53	14.81	0.28	17.50	19.46	1.96
Minnesota	14.01	14.84	0.83	16.74	19.04	2.30
Iowa	14.44	14.52	0.08	17.59	19.28	1.69
Missouri	14.08	14.60	0.52	17.51	19.50	1.98
North Dakota	12.94	13.13	0.19	15.65	17.50	1.85
South Dakota	12.98	13.19	0.22	16.01	17.90	1.90
Nebraska	13.81	14.02	0.21	17.02	18.86	1.84
Kansas	13.96	13.75	-0.21	17.54	18.78	1.24
Delaware	12.91	13.33	0.42	16.78	18.14	1.36
Maryland	13.65	14.15	0.51	16.97	19.28	2.31
Virginia	14.47	14.18	-0.29	17.83	19.32	1.50
West Virginia	14.02	13.96	-0.06	17.60	19.02	1.42
North Carolina	15.08	14.06	-1.02	18.47	19.27	0.80
South Carolina	14.53	14.98	0.44	17.91	20.18	2.26
Georgia	14.51	14.77	0.26	17.95	19.75	1.80
Florida	15.04	13.50	-1.54	18.88	18.99	0.11
Kentucky	14.59	14.78	0.19	18.12	19.89	1.77
Tennessee	14.91	15.28	0.38	18.37	20.44	2.07
Alabama	14.73	14.97	0.24	18.35	20.16	1.81
Mississippi	14.26	14.48	0.22	17.59	19.43	1.84
Arkansas	14.17	14.47	0.30	17.47	19.56	2.09
Louisiana	14.42	13.78	-0.65	17.97	18.75	0.78
Oklahoma	14.46	14.74	0.28	18.41	19.74	1.33
Texas	15.43	16.21	0.78	19.07	21.35	2.29
Montana	13.15	13.41	0.26	16.58	18.31	1.74
Idaho	13.29	13.87	0.58	16.81	18.73	1.92
Wyoming	12.61	12.86	0.25	16.08	17.67	1.59
Colorado	13.78	14.02	0.24	17.27	19.16	1.89
New Mexico	13.63	14.11	0.48	17.37	19.05	1.68
Arizona	14.35	15.29	0.94	18.26	20.41	2.16
Utah	13.67	14.47	0.80	17.24	19.41	2.17
Nevada	13.19	14.46	1.27	16.76	19.30	2.55
Washington	14.21	14.01	-0.20	17.84	18.92	1.07
Oregon	13.77	13.32	-0.45	17.48	18.38	0.90
California	15.22	13.37	-1.85	19.08	18.52	-0.56



(a) Model Land Value in 1979

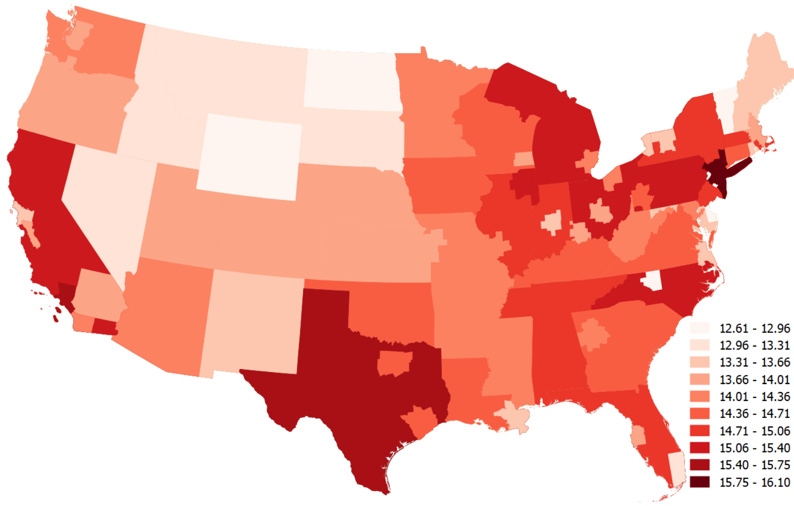


(b) Model Land Value in 2015

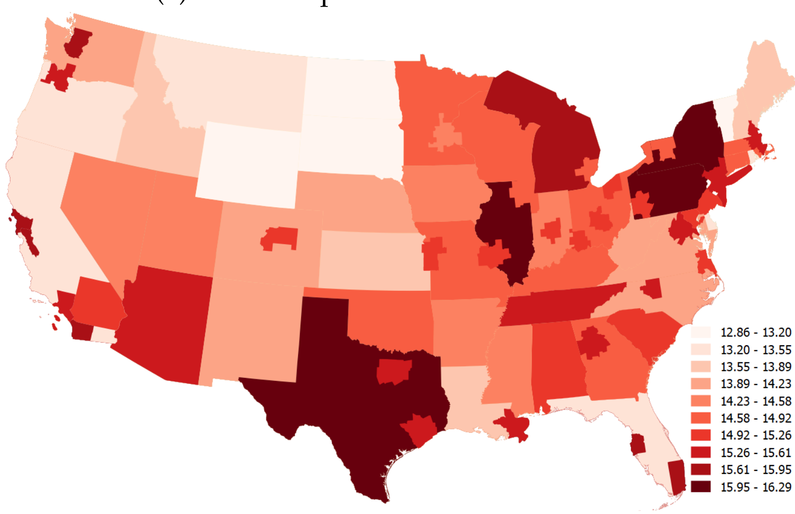


(c) Land Value Change 1979-2015

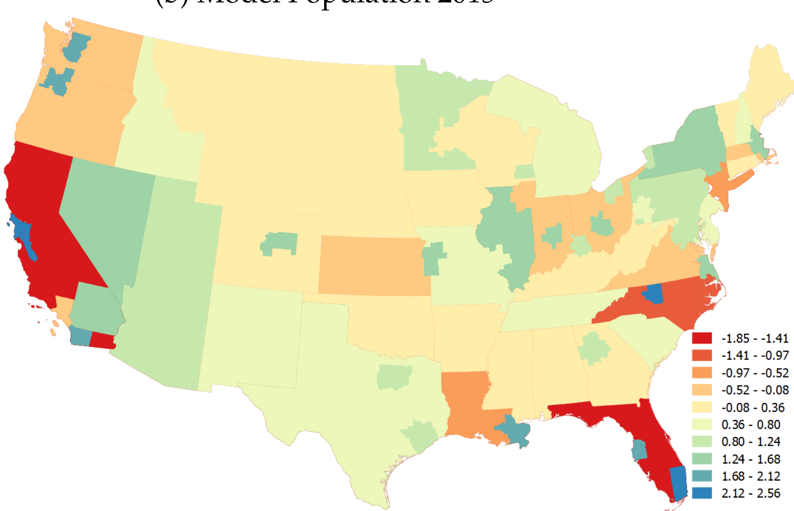
Figure A5: Average Land Value



(a) Model Population 1979



(b) Model Population 2015



(c) Population Change 1979-2015

Figure A6: Average Population and Counterfactuals with Elastic Labour Supply