WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 13-25
CONGESTION, AGGLOMERATION, AND THE STRUCTURE OF CITIES

Jeffrey C. Brinkman
Federal Reserve Bank of Philadelphia

May 23, 2013

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA
Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Congestion, Agglomeration, and the Structure of Cities[1]

Jeffrey C. Brinkman

Federal Reserve Bank of Philadelphia

May 23, 2013

[1]For their comments and suggestions, I thank Nathaniel Baum-Snow, Jan Brueckner, Gerald Carlino, Satyajit Chatterjee, Daniele Coen-Pirani, Dennis Epple, Burcu Eyigungor, Maria Marta Ferreyra, Jeffrey Lin, Antoine Loeper, Tomoya Mori, Jordan Rappaport, Stephen Redding, Esteban Rossi-Hansberg, Holger Sieg, William Strange, and Matthew Turner; seminar participants at Carnegie Mellon University, the University of Pennsylvania, and the Federal Reserve Bank of Philadelphia; and conference participants at CURE, Federal Reserve System Regional Meetings, Midwest Economics Association, NARSC-UEA, and a housing dynamics workshop at Duke University. In addition, I thank Jennifer Knudson for valuable research assistance. The views expressed here are those of the author and do not necessarily represent the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. This paper is available free of charge at www.philadelphiafed.org/research-and-data/publications/working-papers. *Address: Federal Reserve Bank of Philadelphia, 10 Independence Mall, Philadelphia, 19106; Email: Jeffrey dot Brinkman at phil dot frb dot org*

Abstract

Congestion pricing has long been held up by economists as a panacea for the problems associated with ever increasing traffic congestion in urban areas. In addition, the concept has gained traction as a viable solution among planners, policymakers, and the general public. While congestion costs in urban areas are significant and clearly represent a negative externality, economists also recognize the advantages of density in the form of positive agglomeration externalities. The long-run equilibrium outcomes in economies with multiple correlated, but offsetting, externalities have yet to be fully explored in the literature. To this end, I develop a spatial equilibrium model of urban structure that includes both congestion costs and agglomeration externalities. I then estimate the structural parameters of the model by using a computational solution algorithm and match the spatial distribution of employment, population, land use, land rents, and commute times in the data. Policy simulations based on the estimates suggest that naive optimal congestion pricing can lead to net negative economic outcomes.

JEL Classifications: R13, C51, D62, R40

Keywords: Congestion, agglomeration, externalities, spatial equilibrium, urban structure, estimation

1 Introduction

For much of the later half of the 20th century, the focus of transportation planning centered around increasing road capacity in urban areas. While this policy undoubtedly had positive effects by reducing transportation costs and increasing access to land that was previously underutilized, it ultimately led to ever increasing traffic congestion and the realization at the turn of the 21st century that increased capacity was no longer useful in reducing congestion or improving urban mobility. The ineffectiveness of new capacity stems in part from the fact that people and businesses are mobile and can choose where to locate and where to travel within urban areas. In other words, increased capacity can lead to higher travel demand instead of lowering congestion appreciably, which is referred to as induced demand. People respond to changes in transportation costs by changing location and their commuting behavior in cities. Researchers have studied the effects of various transportation policies on mobility, congestion, and urban structure.[1]

Given rising trends, policymakers and planners have become increasingly interested in innovative ways to deal with traffic congestion. These policies range from adding carpool lanes, which is common in the United States, to rationing driving by allowing only certain license plate numbers to enter heavily congested areas on certain days or at certain times, which is common in Latin America. Even transit investment and ridership have shown signs of reversing long declining trends.

Nonetheless, for economists, the holy grail of traffic congestion mitigation has long been congestion pricing or congestion tolls.[2] Congestion in urban areas clearly represents a negative externality in the sense that one person's commuting decision places costs on other commuters. Therefore an efficient policy would be to internalize those costs by taxing commuters at a level equal to the marginal social cost of their commuting decisions. For

[1]Baum-Snow (2007) estimates the effect of highway construction on suburbanization in the 20th century. Duranton and Turner (2011) look at the role of transportation infrastructure and service provision on congestion. A summary of the economics of urban transportation is provided by Small and Verhoef (2007)

[2]See Arnott, DePalma, and Lindsey (1993) for a detailed treatment of congestion and congestion pricing.

policymakers, this has the added benefit of being an additional revenue source. Despite the fact that congestion pricing has been slow to gain public acceptance, recently there have been examples of its implementation, most notably in London and other western European cities.

Congestion pricing would seem an unambiguously positive policy prescription. However, there are also advantages to density. Perhaps the most important advantage of density is the presence of agglomeration externalities, which, in the most basic sense, is simply the fact that productivity improves when firms locate in close proximity to one another. The literature regarding the nature and mechanisms behind agglomeration externalities is well developed[3], and while much of the previous research has focused on agglomerations at a regional scale, there is empirical evidence that agglomeration economies are important even on a neighborhood level. Most notably, Rosenthal and Strange (2003) and Arzaghi and Henderson (2008) find that the production advantages of proximity can decline very rapidly across distances of a few miles or even a few city blocks.

What this indicates is that there are two offsetting externalities at work in urban areas, a negative congestion externality and a positive agglomeration externality, and both are related to the clustering of employment. Therefore, policies designed to deal with either of these externalities in isolation could have unintended or ambiguous consequences, which suggests that better understanding of these two urban spatial mechanisms is warranted. In addition, there might be implications for policies in other types of economies characterized by correlated but offsetting externalites. This framework forms the basis for the analysis in this paper.

The first goal of this paper is to better understand and explain the observed spatial distribution of economic activity in urban areas. To that end, I develop a model of urban

[3]Agglomeration economies were perhaps first proposed by Marshall (1890) and then further dissected by Jacobs (1969). Important work on the scale, structure, and determinants of agglomeration externalities includes, among others, Rosenthal and Strange (2001), Duranton and Puga (2001), Ellison and Glaeser (1997), and Henderson, Kunkoro, and Turner (1995). For a review of the empirical literature see Rosenthal and Strange (2004). For theoretical foundations, see Fujita and Thisse (2002) and Duranton and Puga (2004).

spatial structure that includes both positive agglomeration externalities as well as negative congestion externalities.[4] To date, there has been little research studying the full equilibrium outcomes in the presence of these two offsetting externalities and the consequences for policy, although several theoretical papers have approached the subject. Research by Anas and Kim (1996) perhaps represents the first equilibrium model to include different types of externalities with mobile agents and complete land markets. Later, Arnott (2007), using a simple theoretical set-up, more explicitly made the point that congestion and agglomeration represent offsetting externalities and therefore lead to ambiguous policy prescriptions in urban areas. Ng (2012) further examines the theoretical and policy implications of congestion and agglomeration using a more sophisticated equilibrium model. Finally, this research is related to work by Mayer and Sinai (2003) and Brueckner (2002) who study congestion pricing in the airline industry in the presence of network effects and market power, respectively.

This paper, however, is the first to bring data to a spatial equilibrium model of city structure in order to estimate the magnitude of congestion and agglomeration externalities, along with other important structural parameters related to consumption and production in urban areas.[5] In order to accomplish this, a model is needed that can realistically capture the observed characteristics of a city. I start with the theory introduced by Lucas and Rossi-Hansberg (2002), and then modified and analyzed by Chatterjee and Eyigungor (2012), who develop a circular city framework with transportation costs and agglomeration economies, where firms and workers are free to locate anywhere in the city. I then add an endogenous congestion externality to the model. Finally, I modify the land use specification in order to capture the observed characteristics of land use, where both the extent and intensity of land use are important features of urban structure, and also to capture the observed transitions

[4]Important work on urban spatial structure starts with classic circular city models by Von Thünen (1826), Mills (1967), and others. Work by Fujita and Ogawa (1983) presents an urban spatial model where both firms and workers are mobile. See Anas, Arnott, and Small (1998) for a review of research on urban spatial structure.

[5]The most similar work is by Ahlfedlt, Redding, Sturm, and Wolf (2013), who estimate an equilibrium model that includes transportation costs but not a congestion externality explicitly. Other examples of structural estimation of equilibrium location models with externalities include Holmes (2005), Brinkman, Coen-Pirani, and Sieg (2012), and Davis, Fisher, and Whited (2011).

between commercial, residential, and open space or agricultural land use.

I solve the model using a computational algorithm and then calculate all of the equilibrium quantities and prices in the economy, including those related to endogenous land and labor markets, and endogenous congestion and agglomeration. The solution algorithm implicitly defines a nonlinear mapping from a parameter vector to a set of equilibrium outcomes in the economy. Given this, it is convenient to match equilibrium outcomes from the model to those observed in the real economy, using standard generalized method of moments theory to define the estimator.

The model is estimated using census data, as well as parcel-level land use data, by matching the spatial distribution of employment and residential location, land use patterns, commute times, and land rents. Overall, the model produces a good fit of the underlying characteristics of urban areas. Most notably, the model is able to capture complex land use patterns characterized by the relative level of commercial, residential, and agricultural use across space. Mixing of different land uses occurs at varying levels across space in cities, meaning that the extent of land use is as much an important feature of urban spatial structure as the intensity of land use.

Comparative statics based on the estimates both confirm previous findings about urban spatial structure, and provide new insights. These results show that lower transportation costs lead to dispersed residential location and concentrated employment. On the other hand, if proximity becomes less important for productivity, employment will disperse and residential population will become more concentrated. In addition, the type of transportation infrastructure provided in urban areas, in terms of the relative efficiency in congested versus uncongested areas, is an important consideration along with the level of infrastructure provision.

Finally, the estimated model is particularly well suited to study policies designed to mitigate the negative effects of congestion. Therefore, I simulate a naive optimal congestion pricing policy, where a congestion tax is implemented equal to the marginal social cost of congestion. Revenues are then returned to workers in a non-distorting lump-sum tax.

The results based on point estimates show that congestion pricing actually has a very small net negative effect on all measures of economic welfare — the insight being that congestion pricing leads to more dispersion of employment, and in turn lost productivity, which completely offsets the positive effects from lower congestion costs. This result has very important ramifications for urban policy.

The rest of the paper is organized as follows. Section 2 establishes some empirical regularities of urban structure using data from select cities. Section 3 develops the model, defines equilibrium conditions, and discusses the characteristics of equilibrium. Section 4 describes the algorithm and methods used to solve for equilibrium. Sections 5 outlines the estimation procedure and discusses in detail the variation in the data that identifies structural parameters in the model. Section 6 presents the estimation results and discusses the fit of the model. Sections 7 and 8 present comparative statics and policy simulations based on the estimated parameters of cities. Finally, Section 9 concludes.

2 Evidence on the Structure of Cities

Before introducing the theoretical model, it is important to establish some basic empirical regularities about the structure of cities. In analyzing the data, we want to keep in mind a basic framework for understanding the spatial distribution of activity in urban areas. It has been established that the spatial distribution of economic activity in cities results from competing aggregate forces that can be categorized as contraction forces and expansion forces. Contraction forces lead to concentration and include agglomeration externalities, where firms prefer proximity to other firms, and commuting costs, where individuals prefer to live closer to where they work. Expansion forces primarily are driven by scarcity of land, as well as the demand by both firms and workers for more land and, in turn, cheaper rents. In addition, congestion acts as an expansion force by making dense areas less desirable.

Given this framework, the characteristics of interest for this research lie primarily in the spatial distribution of residential density, commercial density, land use, wages paid, and commute times. In particular, we are interested in how these quantities change in relation to the distance from dense business districts. For illustrative purposes, data are presented for three metropolitan areas: Columbus, Ohio; Philadelphia; and Houston. These cities differ in both size and transportation networks as illustrated in Table 1, which allows for a point of reference when comparing the cities. In particular, Houston and Philadelphia are considerably larger than Columbus, while Philadelphia is the only city of the three with significant transit infrastructure and use.

	Columbus	Houston	Philadelphia
Total Employment (MSA)	845,815	2,100,380	2,559,383
Percent Transit Commuting (MSA)	0.97	3.3	9.2

Table 1: Employment and commuting mode characteristics of selected cities.

Data on employment, residential worker population, wages, and commute times were collected from the 2000 Census Transportation Planning Package available through the Bureau of Transportation Statistics.[6] The data set is a subset of the Decennial Census Sample Data, but is organized geographically both by residential and employment location of workers, to allow for analysis of employment, residential, and commuting patterns in a spatial context. Additionally, the 1992 National Land Cover Database from the U.S. Geological Survey is used to illustrate the relative extent of residential and commercial land at different location.[7] All data were processed using GIS software to calculate all variables as a function of the distance from the central business districts, which are fairly well defined and are dominant employment centers for all three cities. Data were averaged over one-mile increments to smooth the spatial trends for clarity.

Figure 1 shows residential and employment densities for all three cities as a function of distance from the city center. In all three cities, residential densities decline substantially

[6] The data set is available at http://www.transtats.bts.gov.

[7] The data set is available at http://landcover.usgs.gov. More recent land cover data are available from the USGS, but 1992 was the last year that a distinction was made between commercial and residential uses. Note also that the data are constructed using satellite imagery, but supplemented with census data to help distinguish uses in urban areas. This makes the data set useful for initial investigations or for understanding the extent of urbanization as in Burchfield, Overman, Puga, and Turner (2006). However, more detailed sources are preferred for rigorous study of land use types within urban areas. When estimating the structural model, I use parcel-level assessor data.

from the center of the city outward. However, Philadelphia is unique in that it maintains a much higher residential density near the city center. Houston and Columbus follow similar patterns (taking the overall size difference of the cities into account), with little residential density in the central business district, followed by higher density and then gradually declining density.

Employment densities for all three cities are similar to residential densities in that they all decline moving away from the city center. However, employment is much more clustered relative to residential population. In all three cities, there is extremely high employment density at the center, followed by sharp declines, and the gradients are much steeper than residential density gradients. This is true even in Philadelphia, which displays significant residential density. For example, the maximum employment density for a census tract in Philadelphia is over 220,000 employees per square mile, while the maximum residential density is only 35,000 workers per square mile. The discrepancies in Houston and Columbus are even greater.

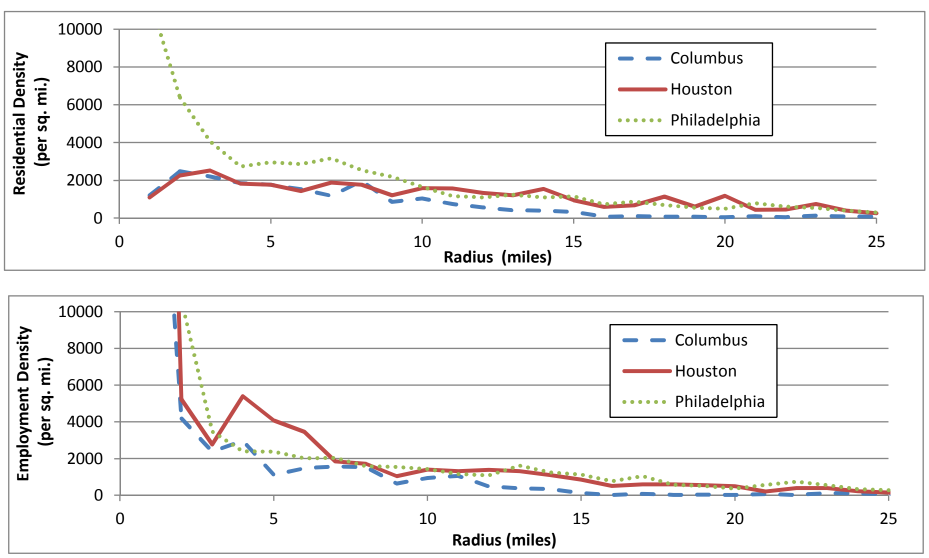

Figure 1: Residential and employment densities as a function of radius from the city center for selected cities. Data source: U.S. Census Transportation Planning Package, 2000

Figure 2 illustrates a different aspect of urban structure, which is the amount of land used for different purposes. A similar pattern exists for all three cities. The percentage of land used for commercial purposes declines rapidly from the center and then continues steadily downward. Residential land use, on the other hand, actually increases initially and then declines steadily. Finally, other uses (mostly agricultural) increase steadily until finally dominating in rural areas. These land use patterns allow us to decompose the densities shown in Figure 1. The densities were calculated by dividing the population by total land area, as is the common way to present densities. However, higher employment density, for example, can result from squeezing more employees into the same amount of land, or by increasing the amount of land used for production in a given location relative to residential land use. The land use plot suggests that the latter is at least as important as the former in characterizing urban structure. In other words, it is important to consider the extensive margin as well as the intensive margin in urban land use.

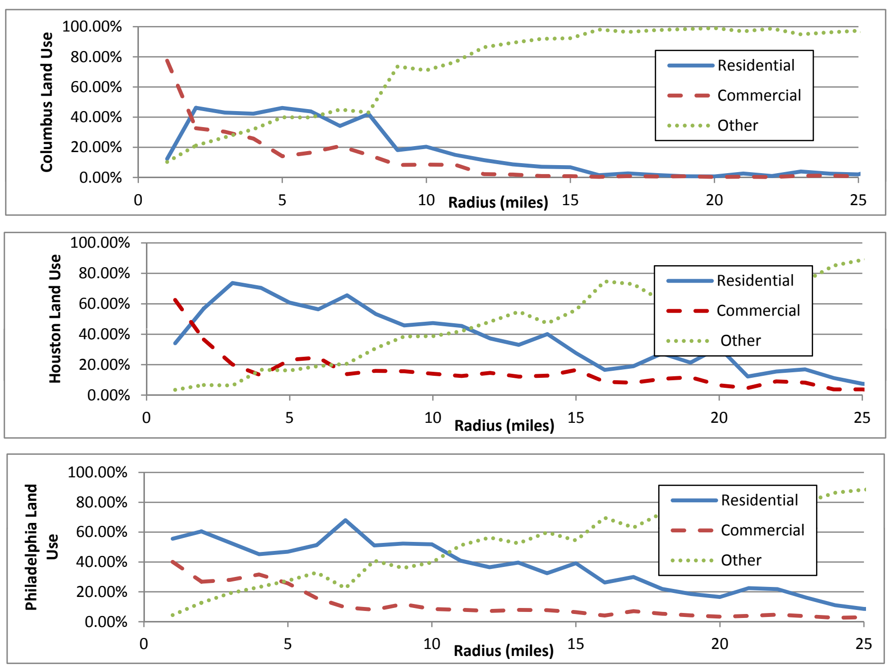

Figure 2: Land use patterns as a function of radius from the city center for selected cities. Data source: USGS National Land Cover Database, 1992

Finally, consider the prices and costs associated with different locations in the cities. Figure 3 shows average one-way commute times by residential location and average annual income by work location.[8] In Columbus and Houston, commute times rise the farther away people live from the city. For Philadelphia, commute times rise, then decline, and then flatten.[9] The wages paid by firms, on the other hand, decline away from the center, with a particular wage premium attached to the central business district. Given heterogeneity in both firms and workers, the raw wage data should be interpreted with a measure of skepticism. However, it seems clear that there is a negative correlation between commute times and wages across space. [10]

Overall, the data suggest that agents in the economy are faced with trade-offs associated with their location decision. First, given that there appears to be a wage premium in dense areas, firms must be willing to pay these prices to take advantage of production gains in dense business districts. Second, it appears that workers will sacrifice land consumption for shorter commute times. These effects are exhibited in densities, land use patterns, wages, and commute times for all three cities. Third, jobs are much more clustered than residents, which suggests that different spatial forces work on firms versus workers. In general, the goal of this research is to explain the characteristics of these data, by modeling and estimating the trade-offs and equilibrium effects of agglomeration and commuting costs, while paying special attention to the effects of commuting congestion.

[8] Land rents are also clearly an important price consideration in urban city structure. However, rent data are not available for all three cities. Later, land values are shown for Columbus and used in the estimation of the structural model. Not surprisingly, land prices decline away from the city center, with commercial prices declining at a faster rate than residential prices.

[9] Commuting patterns in the Philadelphia area are complex, given that there are multiple employment centers close to Philadelphia. The other two cities, however, are more isolated from other major labor markets. The distribution of transit provision and use also skews commute times somewhat.

[10] This is not a new finding. See Timothy and Wheaton (2001) for a more careful examination of this correlation.

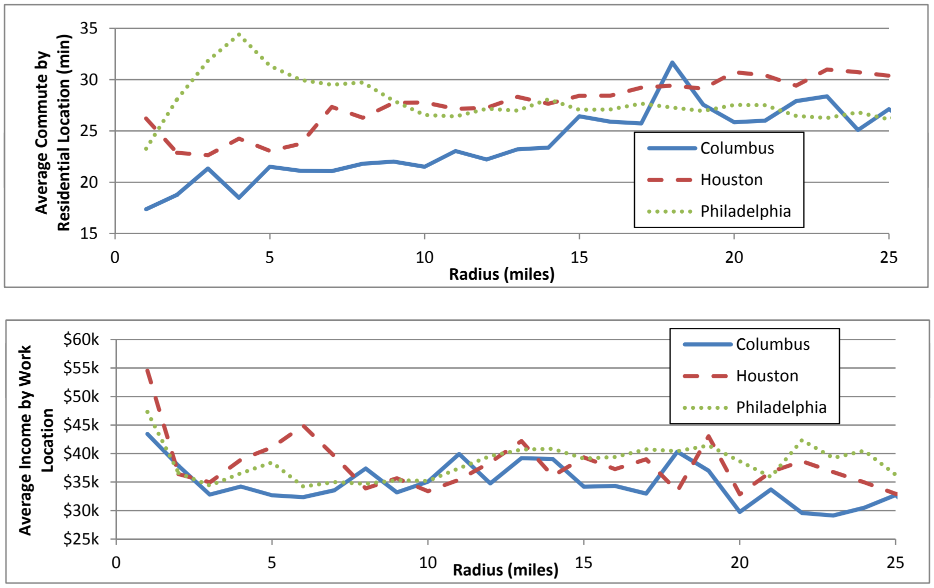

Figure 3: Wages paid at work locations and average one-way commute times by residential location for selected cities. Data source: U.S. Census Transportation Planning Package, 2000

3 The Model

This section develops a model that can capture the important features of an urban economy described above. The model assumes a circular city and considers only symmetric allocations. Beyond these assumptions, no restrictions are placed on the location of businesses or residents. The model draws heavily on the theory provided by Lucas and Rossi-Hansberg (2002) and Chatterjee and Eyigungor (2012). In these papers, the authors discuss the theoretical characteristics of an economy with transportation costs and production externalities. These papers provide a general well-formed theory of the forces affecting job and residential locations within an urban area.

One key theoretical departure of the current research is that congestion is explicitly added to the transportation cost, as opposed to a simple distance cost. Note that this adds a second externality in the model, in that individual workers' commuting decisions now place an external cost on the economy. By more precisely modeling the transportation costs in the economy, we can study policies designed to remedy congestion in urban areas. Furthermore, this modification allows for transportation technologies that have variable costs in different levels of congestion and therefore different effects on the spatial structure of the city in equilibrium.

In addition, the model has been modified to allow for more mixing of commercial and residential uses. Previous models place sharp restrictions on land use and allow for mixing of uses only under precise conditions. Empirical evidence suggests, however, that there is a great deal of mixing of land uses, and the extent of land used for different purposes is an important feature of urban spatial structure. Furthermore, observed land use patterns exhibit gradual transitions in the data. To address this disconnect between the theory and data, a local complementarity of uses is added to the model. These changes, along with all the assumptions of the theoretical model, are discussed here. The model, as formulated, is able to produce a wide range of employment and residential distributions and commuting patterns. A precise description of the model follows.

Assumption 1 *The city is circular and allocations are symmetric. Additionally, commuting is always radial, and the size of the city is fixed at a radius, S.*

The assumption of symmetric equilibria implies that all allocations can be written as a function of r, defined as the distance from the center of the city, and not the angle, ϕ (using standard circular coordinates). The assumption of circular symmetry is strong; however, these symmetric equilibria are feasible in a general sense. Still, symmetric equilibria are not necessarily stable in the presence of asymmetric shocks. Nonetheless, in the interest of tractability and generalizability, the circular assumption is useful.

At this point some additional notation and definitions will make description of the model clearer. $\theta_C(r)$ is the fraction of land used for production/commercial use at location r, $\theta_R(r)$ is the fraction of land used for consumption/residential use, $\theta_A(r)$ is the fraction of open space or agricultural land, and $\theta_C(r)+\theta_R(r)+\theta_A(r) = 1$. $n_C(r)$ is the employment intensity at r defined as employment per unit of commercial land[11], and $n_R(r)$ is the residential intensity defined as the number of workers housed per unit of residential land. In addition, $q_C(r)$, $q_R(r)$, and q_A are commercial bid rent, residential bid rent, and agricultural bid rent, respectively.[12] Given these definitions, we can now describe the preferences of consumers and the technology of producers.

Assumption 2 *Workers supply one unit of labor inelastically and have increasing preferences over general consumption, $c(r)$, and land, $l(r)$.*

Worker preferences are modeled using a Cobb-Douglas form, given by $U(c(r), l(r)) = c(r)^{\beta} l(r)^{1-\beta}$. Given this functional form, β is interpreted as a consumption share parameter, or the share of expenditures on all goods except land.

Assumption 3 *Firms produce a good to be sold in an external market through a production function that is increasing in land and labor, and firms are subject to an agglomeration externality.*

The production function is modeled as constant returns and Cobb-Douglas. The production per unit of land is given by

$$x(r) = g(z(r))f(n_C(r)),$$

where g is the production externality, given by

$$g(z(r)) = z(r)^{\gamma},$$

[11] Note that "intensity" is used to distinguish from the standard definition of density. Intensity here is defined as employees per unit of commercial land as opposed to employment per unit of total land, which is the standard definition of density. Intensity describes the reciprocal of land demand per employee by firms. Density can be calculated by the product of intensity and commercial land use, $\theta_C(r)n_C(r)$. Residential intensity is analogous.

[12] Agricultural rent is assumed to be independent of location and exogenously given.

and f describes the land and labor technology, given by

$$f(n_C(r)) = An_C(r)^{\alpha}.$$

In the above form, γ determines the scale of the externality, while α is the ratio of the share of labor and land. Also, note that the constant returns specification allows for a single production input variable defined as labor per unit of land, as opposed to including land and labor separately.

Assumption 4 *The externality depends on a proximity measure, $z(r)$, of a firm located at $(r, 0)$ to other firms in the economy located at (s, ϕ). This measure is assumed to be increasing in employment of other firms and attenuates with distance from other firms.*

For the agglomeration externality, I adopt the specification of Chatterjee and Eyigungor (2012), who assume that the production externality travels only along rays.[13] With this assumption, the relevant distance between a firm located at radius r and a firm located at radius s is just $(r + s)$. This simplifies the model considerably by ensuring an exponential decay of the agglomeration externality with distance. The externality at a location r is then given by

$$z(r) = \int_0^S 2\pi s \theta_C(s) n_C(s) e^{-\delta(r+s)} ds, \tag{1}$$

where δ determines the rate of exponential decay of the externality with distance. This leads to a very simple functional form of the externality, for a known productivity at the center of the city, $z(0)$, given by

[13]Using this specification, the authors can solve for the equilibrium analytically and show that, for a fixed city size, equilibrium is unique. For the current research, this specification has several advantages over the externality specification of Lucas and Rossi-Hansberg (2002) who assume that the externality travels in all directions in space. First, computation becomes simpler, by avoiding numerical computation of a two-dimensional integral. Second, even given the additional complexity of the current model, there is strong evidence of uniqueness of equilibrium. (Uniqueness is discussed in more detail later in the paper and in Appendix A.) Uniqueness and computation speed are both very important for structural estimation, given that I use a fixed-point algorithm nested inside a parameter search routine to estimate the model. Finally, and most importantly, this specification seems to provide a better fit for the observed gradients in the data.

$$z(r) = z(0)e^{-\delta(r)}, \tag{2}$$

where

$$z(0) = \int_0^S 2\pi s \theta_C(s) n_C(s) e^{-\delta s} ds. \tag{3}$$

The next component of the economy is a commuting cost.

Assumption 5 *Workers pay a cost to travel to work that is subtracted from their wage. This cost is increasing both in distance traveled and in congestion.*

The commuting cost is paid by workers, meaning that each worker always delivers one unit of labor to the location of the firm but experiences a loss of wages equal to the commuting cost.[14] This means that the wage that enters into the worker's budget constraint is the wage she earns minus commuting costs. Furthermore, the current model assumes that this commuting cost is a function of both distance and congestion. The total commuting cost between a location s and location r will be denoted $T(s,r)$ and can be written using the following parameterization for the marginal cost of commuting through location r as

$$T'(r) = w(r)(\tau + \kappa M(r)) \tag{4}$$

where $M(r)$ is the mass of commuters traveling through a location r and is a measure of congestion. Note that the marginal cost of traveling through r is dependent on the wage at r, meaning that the commuting costs are modeled as a percentage of lost wages. The parameters τ and κ can be interpreted as the distance and congestion costs, respectively. Using this specification, we can consider transportation technologies (or mixes of technolo-

[14] Lucas and Rossi-Hansberg (2002) model the commuting cost as a loss of labor to firms. This difference in specification does not change the incidence of transportation cost in equilibrium, only the interpretation of who pays the cost. The former specification has advantages in theoretical analysis, while the current specification enhances empirical analysis by allowing the researcher to relate model characteristics more directly to data such as residential location, job location, and commute times.

gies) that exhibit different costs in the presence of different congestion levels. For example, in the context of intraurban transportation, one could conjecture that transit systems provide relatively low costs in congested areas, while automobiles provide relatively low costs in uncongested areas. This parameterization also isolates the effect of congestion on commuting cost. These costs do not include the public costs of construction, maintenance, or operation of transportation services, but merely the costs faced by individual commuters.

Given the dependence of transportation costs on congestion, it is important to further investigate the congestion function, $M(r)$. Consider that the total number of commuters at a location, r, is the sum of all jobs from 0 to r less the number of residents from 0 to r. The congestion function is the following:

$$M(r) = \int_0^r 2\pi r'[n_C(r')\theta_C(r') - n_R(r')(\theta_R(r')]dr'. \tag{5}$$

This congestion function can be positive or negative, where positive values represent commuting toward the center of the city, while negative values represent commuting away from the center. A value of zero represents zero commuting and also represents a radius containing equal masses of jobs and residents. Note that $M(r)$ is determined directly from $n_C(r)$, $n_R(r)$, $\theta_C(r)$, and $\theta_R(r)$.

In terms of land use, previous models often assumed that land is allocated to the highest bidder (i.e., if the commercial bid rent is higher than the residential bid rent, then all land is used for commercial purposes and vice versa). This sharp restriction is hard to justify empirically given that, in reality, one observes a large amount of mixing of uses at all geographic scales. This points to some complementarity of uses at a very local neighborhood level. While I do not specifically study the determinants of this mixing, in order to successfully compare the model to the data, I include a local complementarity of uses in the specification of land supply.[15]

[15] Other model specifications could produce similar land use functions. For example, idiosyncratic location-specific preference or production shocks at the individual or firm level could lead to probabilistic specifications for land use that would take a similar form. See Anas and Kim (1996) for an example.

Assumption 6 *Land is owned by an absentee landlord. The landlord will seek to maximize rent revenue per unit of land at every location, but is subject to a transformation function for land services describing the complementarity of land uses.*

A constant returns, constant elasticity function describes this transformation of land services. The landlord's maximization problem for commercial, residential, and agricultural allocation, $\theta_C(r)$, $\theta_R(r)$, and $\theta_A(r)$, respectively, at each location r, for commercial, residential, and agricultural rents, $q_C(r)$, $q_R(r)$, and q_A respectively, is given by,

$$\max_{\theta_R(r),\theta_C(r),\theta_A(r)} \quad q_R(r)\theta_R(r) + q_C(r)\theta_C(r) + q_A\theta_A(r) \tag{6}$$

$$s.t. \quad \left[\eta_R\theta_R(r)^{\frac{\rho-1}{\rho}} + \eta_C\theta_C(r)^{\frac{\rho-1}{\rho}} + \eta_A\theta_A(r)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}} = C,$$

where η_R, η_C, and η_A are share parameters summing to 1, $\rho \in [0,\infty]$ is the elasticity between land uses, and C is some scale constant.[16]

The equilibrium definition invokes some of the usual conditions. Firms will maximize profits at every location, r, given wages, $w(r)$, commercial rents, $q_C(r)$, and the productivity, $z(r)$. Workers will maximize utility at every location, given net wages, $w(r)$, and residential rents, $q_R(r)$. The landlord will maximize rent at every location given commercial rents, $q_C(r)$, and residential rents, $q_R(r)$. In addition, the market for land must clear at every location.

However, several other conditions are needed to describe equilibrium in a spatial context, given the mobility of both firms and workers.

Assumption 7 *The city exists in a larger economy, and firms and workers alike are free to enter or exit.*

The above assumption implies a zero profit restriction on firms in equilibrium. For work-

[16]This specification produces a well-behaved, continuous land use function that simplifies equilibrium computation and is also consistent with observed land use patterns.

ers, the assumption suggests that there is a reservation utility, $\bar{u}$, which must be achieved with equality at every location, r, in the city in equilibrium. Since workers and firms will achieve identical utility and profits, respectively, at every location, then they will have no incentive to relocate in equilibrium; therefore, no additional condition is necessary to ensure that firms or workers cannot be made better off by relocating within the city.

However, workers not only decide where to live; they also are free to choose where they work. To ensure that workers have no incentive to commute to a different location, equilibrium requires a restriction on the wage gradient through commuting costs. The necessary condition in equilibrium is that the difference in wages paid between two locations must be equal to the total commuting cost of traveling between the two locations. Given the functional form of the marginal commuting cost, this condition can be written as

$$w(r) - w(s) \leq \int_s^r w(r)(\tau + \kappa M(r'))dr', \forall r, s \in [0, S]. \tag{7}$$

It is straightforward that workers will only travel toward wages that are higher than the wages paid where they live. To gain further intuition into this condition, consider the situation where the difference in wages is greater than the commuting cost between locations. If this were the case, then workers would all desire to work at the high-wage location. Likewise, if the difference in wages were less than commuting costs between locations, then all workers would desire to work at the low-wage location. Given that the land use function requires some employment at all locations, this condition must hold with equality everywhere. Note that, given the dependence on congestion, $M(r)$, the wage path cannot be written as a simple exponential decay function of radius, r, as is possible with the productivity function, $z(r)$. This is an important source of nonlinearity in the model.

Finally, closing the model requires a labor-market-clearing condition, which states that all workers must be housed within the city. An equivalent condition is that commuting is zero at the edge of the city, which can be formally written as $M(S) = 0$. The underlying assumption here is that commuting costs from other cities are sufficiently high to prevent such activity. All of the pieces are now in place to formally define equilibrium.

Definition 1 *Equilibrium is defined as a set of allocation functions* $\{z(r), \theta_R(r), \theta_C(r), n_C(r), n_R(r), M(r)\}$ *along with a set of price functions* $\{w(r), q_C(r), q_R(r)\}$ *defined on [0,S], such that for all r:*

i . Firms choose $n_C(r)$ *to maximize profits at all locations, given* $z(r)$*,* $w(r)$*, and* $q_C(r)$*.*

ii . Workers choose $n_R(r)$ *to maximize utility at all locations, given w(r) and* $q_R(r)$*, subject to their budget constraint.*

iii . Landlords maximize rents given, $q_r(r)$ *and* $q_R(r)$*, subject to equation (6).*

iv . $M(r)$*,* $n_C(r)$*,* $n_R(r)$*,* $\theta_C(r)$*, and* $\theta_R(r)$ *satisfy equation (5), and* $M(0) = 0$*.*

v . Firms achieve zero profits and workers achieve a reservation utility, $\bar{u}$*, at every location, r.*

vi . $w(r)$ *must satisfy equation (7) and the labor market clears, i.e.,* $M(S) = 0$*.*

vii . $z(r)$*,* $\theta_C(r)$*, and* $n_C(r)$ *satisfy equation(1).*

At this point some discussion of the characteristics of equilibrium is warranted. Existence and uniqueness are important given that a computational solution algorithm is used to estimate the structural model. Existence and uniqueness are features of the framework used by Chatterjee and Eyigungor (2012) for a fixed city size. However, given the addition of endogenous transportation costs, and a modified land use function, uniqueness is not as straightforward in the current set-up. Generally speaking, in these types of agglomeration models, if the mapping from $z(r)$ to itself is characterized by a unique and stable fixed point, then there is a unique and stable equilibrium. I am able to verify local uniqueness computationally. Although a full proof remains elusive, there is very strong evidence that this particular set-up exhibits global uniqueness. A partial proof and a more detailed discussion of the uniqueness of equilibrium are found in Appendix A.

4 Equilibrium Solution

This section outlines the computational methods used to calculate equilibrium given the functional form choices introduced above. The computational solution employs a shooting algorithm to find a wage path consistent with the labor-market-clearing condition, $M(S) = 0$, which is nested inside a fixed-point algorithm, ensuring that the agglomeration externality is consistent with the spatial allocation of employment, as in equation (1).

First, note that for a given wage path, $w(r)$, and productivity function, $z(r)$, all other equilibrium allocations and prices can be solved analytically. This allows us to write the equilibrium choices for firms, workers, and landlords as a function of those two variables. For the most part, these allocations and prices are log-linear functions of wages and productivity. It is convenient to write the functions in log form for discussion of identification later on.

Solving the profit maximization problem of the firm, the indirect labor choice per unit of land as a function of wages and the externality is

$$ln\, n_C(r) = \frac{1}{1-\alpha}\left(ln\alpha + lnA - ln\, w(r) + \gamma\, ln\, z(r)\right). \tag{8}$$

Using the indirect labor function, along with the zero-profit assumption, we can solve the equilibrium commercial bid rent, $q_C(r)$, as a function of productivity $z(r)$ and wage $w(r)$:

$$ln\, q_C(r) = \frac{1}{1-\alpha}\left((1-\alpha)ln(1-\alpha) + \alpha\, ln\, \alpha + lnA - \alpha\, ln\, w(r) + \gamma\, ln\, z(r)\right). \tag{9}$$

In a similar fashion, we can solve for an individual's indirect labor intensity choice and equilibrium residential rent. Recall that workers supply one unit of labor, so this is really the reciprocal of land demand per worker. Individuals will maximize utility subject to their wage, $w(r)$. Recall that wage, $w(r)$, takes a different meaning for individuals, in the sense that $w(r)$ is the net wage defined as wage paid less commuting costs. The city's existence in a larger economy suggests that individuals have a reservation utility, $\bar{u}$, which gives the following condition:

$$c(r)^{\beta} l(r)^{1-\beta} = \bar{u}.$$

With this condition, solving for the indirect residential intensity gives

$$ln\, n_R(r) = \frac{1}{1-\beta} \left(\beta\, ln\, \beta - ln\, \bar{u} + \beta\, ln\, w(r)\right) \tag{10}$$

and solving the equilibrium residential bid rent gives

$$ln\, q_R(r) = \frac{1}{1-\beta} \left((1-\beta)\, ln\, (1-\beta) + \beta\, ln\, \beta - ln\, \bar{u} + ln\, w(r)\right). \tag{11}$$

The rent maximization problem of the landlord, along with the transformation of land services, gives the following land supply function which is based on the ratio of commercial and residential rents in a location for commercial land use,

$$\theta_C(r) = \frac{1}{1 + \left(\frac{q_R(r)}{q_C(r)} \frac{\eta_C}{\eta_R}\right)^{\rho} + \left(\frac{q_F}{q_C(r)} \frac{\eta_C}{\eta_F}\right)^{\rho}}, \tag{12}$$

and for residential land use,

$$\theta_R(r) = \frac{1}{1 + \left(\frac{q_C(r)}{q_R(r)} \frac{\eta_R}{\eta_C}\right)^{\rho} + \left(\frac{q_F}{q_R(r)} \frac{\eta_R}{\eta_F}\right)^{\rho}}. \tag{13}$$

This is a very flexible form and allows for complete segregation of uses (corresponding to $\rho = \infty$) or complete mixing of uses (corresponding to $\rho = 0$). In addition, there can be differences in the shares of different uses. An additional benefit of this form is that it produces continuous land use functions, $\theta_C(r)$ and $\theta_R(r)$, which is useful in numerical computation of equilibria for the model.[17]

Finally, the transportation costs, in concert with free mobility of labor, imply that wages

[17] This functional form for land use is similar to that used by Wheaton (2004).

must adhere to the following spatial restriction:

$$w(r) - w(s) = \int_s^r w(r)(\tau + \kappa M(r'))dr', \forall r, s \in [0, S].$$

Notice that for a given congestion, $M(r)$, and an initial wage at the center of the city, $w(0)$, the entire wage path can be calculated. This wage path can be increasing or decreasing, depending on the direction of commuting.

We can now describe the equilibrium solution algorithm. The solution algorithm can be thought of as an inner loop that searches for an initial wage, $w(0)$, leading to a wage path and allocations consistent with equilibrium (conditions $i - vi$) for a given productivity, $z(r)$, and an outer loop that consists of a iterative fixed-point algorithm to find a productivity function consistent with the externality specification(condition vii).

The algorithm starts by guessing an initial productivity function, $z(r)$, which given the form of the externality is completely determined by $z(0)$. In the inner loop of the algorithm, this productivity function is taken as given. The next step is to guess an initial wage, $w(0)$. With the initial wage, we can construct the entire wage path and hence all the allocations of the economy. Given that $w(0)$ and $z(0)$ are known, we can calculate $n_C(0)$ and $n_R(0)$ and the congestion $M(0)$. Knowing the congestion allows for the calculation of the wage at a small distance from the center, $r = \epsilon$. We can then calculate the allocations and congestion at $r = \epsilon$. The algorithm continues to move outward from the city center, until it reaches the edge of the city, $r = S$. At this point, the algorithm checks that the labor market clears, or that $m(S) = 0$. Given the current functional specification, $M(S)$ is a decreasing continuous function of $w(0)$.[18] Therefore, from this point, any minimization routine can find the initial wage such that the labor market clears.

The outer loop of the algorithm then uses the commercial density function, $n_C(r)$, along

[18]Note that the continuity of this function arises from the land supply specification. This reduces the complexity of the solution algorithm used by Lucas and Rossi-Hansberg (2002), where the correspondence between $M(S)$ and $w(0)$ was decreasing but not continuous, given that the land use function was discontinuous. More discussion of the relationship between $M(S)$ and $w(0)$ is found in the discussion of existence and uniqueness in Appendix A.

with the land use function, $\theta_C(r)$, to calculate the theoretical productivity at the center of the city, $z(0)$, and hence the entire productivity function, $z(r)$, using equations (2) and (3). The productivity function is updated, and the routine repeats the inner loop. It continues this process until $z(0)$ converges to a fixed point. At this point the algorithm has found allocations and prices consistent with conditions $i - vii$ of the equilibrium definition.

5 Estimation

The above solution algorithm implicitly defines a nonlinear mapping from the parameter vector of the structural model, from here on denoted as Θ, to a spatial distribution of equilibrium outcomes. This suggests an estimation strategy that matches computational outcomes to observed outcomes in the data. This can be accomplished using standard generalized method of moments (GMM) estimation, as suggested by Hansen (1982). The general methodology is as follows.

Denote the vector of sample moments as m_N, and denote with $\hat{m}(\Theta)$ the vector of equivalent computational moments calculated through the solution algorithm. Using this notation, a vector of orthogonality conditions is defined as

$$g_N(\Theta) = m_N \; - \; \hat{m}(\Theta).$$

Θ can then be estimated using the following consistent estimator, which minimizes the weighted distance between the sample moments and the computational moments,

$$\hat{\Theta}_N = \arg\min_{\Theta} g_N(\Theta)' A_N g_N(\Theta),$$

for some positive semi-definite matrix A_N, which converges in probability to A_0. The asymptotic distribution of the estimator is then given by

$$N^{1/2} \, (\hat{\Theta}_N - \Theta_0) \xrightarrow{d} \; N(0, \; (D_0' \, V^{-1} \, D_0)^{-1}).$$

where V^{-1} is the efficient weighting matrix given by the inverse covariance matrix of $g_N(\Theta)$ and D_0 is the Jacobian of the expected value of $g_N(\Theta)$ with respect to Θ. The asymptotics are all computed numerically. In addition, the minimization of the objective function routine was repeated until the optimal weighting matrix converged. Results appear robust to numerical approximations, including tolerances, discretization, and interpolation.

5.1 Identification

While the methodology above is straightforward, it is important to further discuss identification of the parameter vector and construction of moment conditions. Ultimately, identification of all parameters relies on matching all observed moments simultaneously. The actual estimation of the structural parameter vector requires a computational search, using a functionalized version of the computational equilibrium solution algorithm described earlier. However, the following description should provide intuition and a solid foundation for understanding the sources of variation in the data that identify specific structural parameters in the model, as well as outline the specific moments used in the estimation. None of this relies on using observed wages. Instead, given that we are trying to specifically understand commuting and congestion costs, I use commute times to directly identify commuting costs and hence wages through the wage gradient restriction. The following is a step-by-step description of the identification strategy and the moment conditions used in the estimation.

Assume initially that wages, $w(r)$, and location specific productivity, $z(r)$, are observed. This makes identification of preference, production, and land use parameters straightforward. In fact, the equations mostly take the form of log-linear functions of $z(r)$ and/or $w(r)$, so it is reasonable to use a nonlinear least squares approach as a guide for construction of the GMM moment conditions.[19] The production, preference, and land use are identified as follows.

[19] As in a nonlinear least square approach, we can match the sample means to the computational predictions and assume that errors are uncorrelated with changes in wages, $w(r)$ or productivity $z(r)$. In other words, we can match the slope and intercept of the functions. Given the Cobb-Douglas functional form assumptions, it is reasonable to match the logs of the observed variables, rather than levels.

1. **Preferences:** $\bar{u}$ and β are identified by the level and change of residential intensity, $n_R(r)$, which is observed (See equation 10). The observed residential rents, $q_R(r)$, can be used as an overidentifying restriction (See equation 11)

2. **Production:** A, α, and γ are jointly identified by the level and change of commercial intensity, $n_C(r)$, and the level and change in commercial rent, $q_C(r)$, which are observed (See equations 8 and 9). This system is slightly overidentified.

3. **Land use:** ρ, η_R, and η_C are jointly identified by the residential and commercial land use functions, $\theta_R(r)$ and $\theta_C(r)$, along with the open space rents (i.e. nonurban value of land), $q_{F,}$ which are observed. (See equations 12 and 13.) [20]

Now we can return to the wage gradient, $w(r)$, which was assumed to be known in the argument above. Instead of trying to measure wages directly, we can identify the wage gradient using the observed commute times, given the structure of the model.[21] The commuting cost through a location is determined by the distance traveled, r, and the mass of commuters, $M(r)$, traveling through a location, for given parameters. Therefore, the distance, r, and the level and change of the commuting cost by residential location, $c_t(r)$, along with the observed net commuting patterns, $M(r)$, jointly determine the transportation parameters, τ and κ. Note that this is the first time that spatial relationships, through the radius, r, have been used explicitly in the identification argument.

What this all suggests, in practice, is to match the level and the slope of all variables, including commute times. However, the slope of the gradients depends not only on the distance from the center of the city, but also on the level of congestion. Therefore, we can construct moment conditions by assuming that errors (the distance between observed and computational moments) are uncorrelated with both r and $M(r)$. This is equivalent to a non-linear least squares approach, where the slopes of the functions are determined by a linear combination of observables.

[20] The commercial and residential rents also enter here, but were already used in the identification argument for preference and production parameters; hence the need for land use data.

[21] The average wage for the entire city is used as a moment to identify the level of wages, but commute times are used to identify the slope.

The only consideration left is how to identify the productivity function $z(r)$ and the associated attenuation parameter δ. Obviously, there is no observational equivalent to $z(r)$ in the data, so δ cannot be identified from magnitudes in the data separately from the productivity scale parameter A.[22] Technically, if all the other parameters are known, then δ could be identified from the commercial gradients in the data. However, the slope of commercial gradients cannot be used to identify δ separately from γ given that they both enter equations 8 and 9 in the same way, given the functional form of the externality. This is the technical downside of assuming that agglomeration externalities move only along rays. However, the product of δ and γ can be identified using the commercial gradients. This is enough to fully characterize the change in productivity across space in the current set-up. Therefore, I simply estimate the product, $\gamma * \delta$.[23] This concludes a complete description of the identification strategy for the structural parameters of the model. The precise moment conditions used in estimation are found in Appendix B.

5.2 Data and Implementation

I estimate the model using data from Columbus, Ohio. Columbus is used for several reasons. First, the geography of Columbus matches the assumptions of the model. Columbus has a clear center, along with a radial and uniform automobile-oriented transportation infrastructure. Moreover, it resembles a featureless plain in the sense that there are no geological features like oceans, rivers, or mountains to distort development. In addition, it represents a self-contained labor market, mostly separated from other metropolitan areas and political boundaries. Figure 4 shows the spatial distribution of activity in Columbus. It clearly displays radial, if not circular patterns for many characteristics of interest. It should also be noted that Columbus represents a median U.S. city on dimensions such as demographics,

[22]Note that there would presumably be some scale parameter on the externality specification. This also is not identified separately from A. In other words, only one scale (intercept) parameter can be identified within the production function. In implementation of the estimator, it is convient to normalize $z(0)$ to 1, and then estimate A. Scale parameters have little real-world meaning in this set-up, but are necessary to match the data. The change in values accross space is the focus of this estimation.

[23]In practice, it is convienient to fix γ and estimate δ. This is an equivalent procedure and will not change estimates of any of the parameters of interest.

population, incomes, and education levels. Finally, parcel-level data is available for Franklin County, the central county of the Columbus area, which has information on land use and land prices at a fine geographic resolution. This sort of detailed data is not readily available for all cities.

Based on the discussion of identification above, estimating the model requires data on employment and residential intensity, residential and commercial land prices, commuting costs, and land use patterns. The data are drawn from two sources. The first is the 2000 Bureau of Transportation Statistics's Transportation Planning Package, which was described earlier and has information at the census tract level on employment, residential worker population, and commute times. The second source is year 2000 assessment data from the Franklin County Auditor's office, which contains parcel-level information on assessed land values and land use. This data are aggregated to the tract level for consistency.

Commercial and residential intensity, $n_C(r)$ and $n_R(r)$ are calculated by dividing employment by total commercial or residential land, respectively. Rents $q_C(r)$, $q_R(r)$, and q_A are calculated from the assessed land values and converted to rents using a 5 percent rent-to-value ratio.[24] Commercial and residential land use, $\theta_C(r)$ and $\theta_R(r)$, are calculated as the percentage of each type of use in a census tract from the assessor data.[25] Finally, the commuting cost is calculated from commute times, and converted to commuting costs by using a time-cost conversion factor of half the average wage rate.[26] Technically speaking, the commuting time is observed by residential location and not as a cost of traveling through a location as it is in the model. This discrepancy can be overcome with some additional computation, the technical details of which can be found in Appendix C.

Given that the land use and price data are available only for the county, the sample moments are calculated for the area inside a 12.5-mile radius from the center of the city, which captures a good part of the population, employment, and spatial variation in the

[24] This conversion ratio is not crucial. It will affect the level estimates but not the estimates of relative changes across space.

[25] Commercial land use includes industrial and commercial classifications.

[26] This is a common conversion factor in both the literature and in practice. See Tse and Chan (2003) and Small (1983) for estimates in the literature.

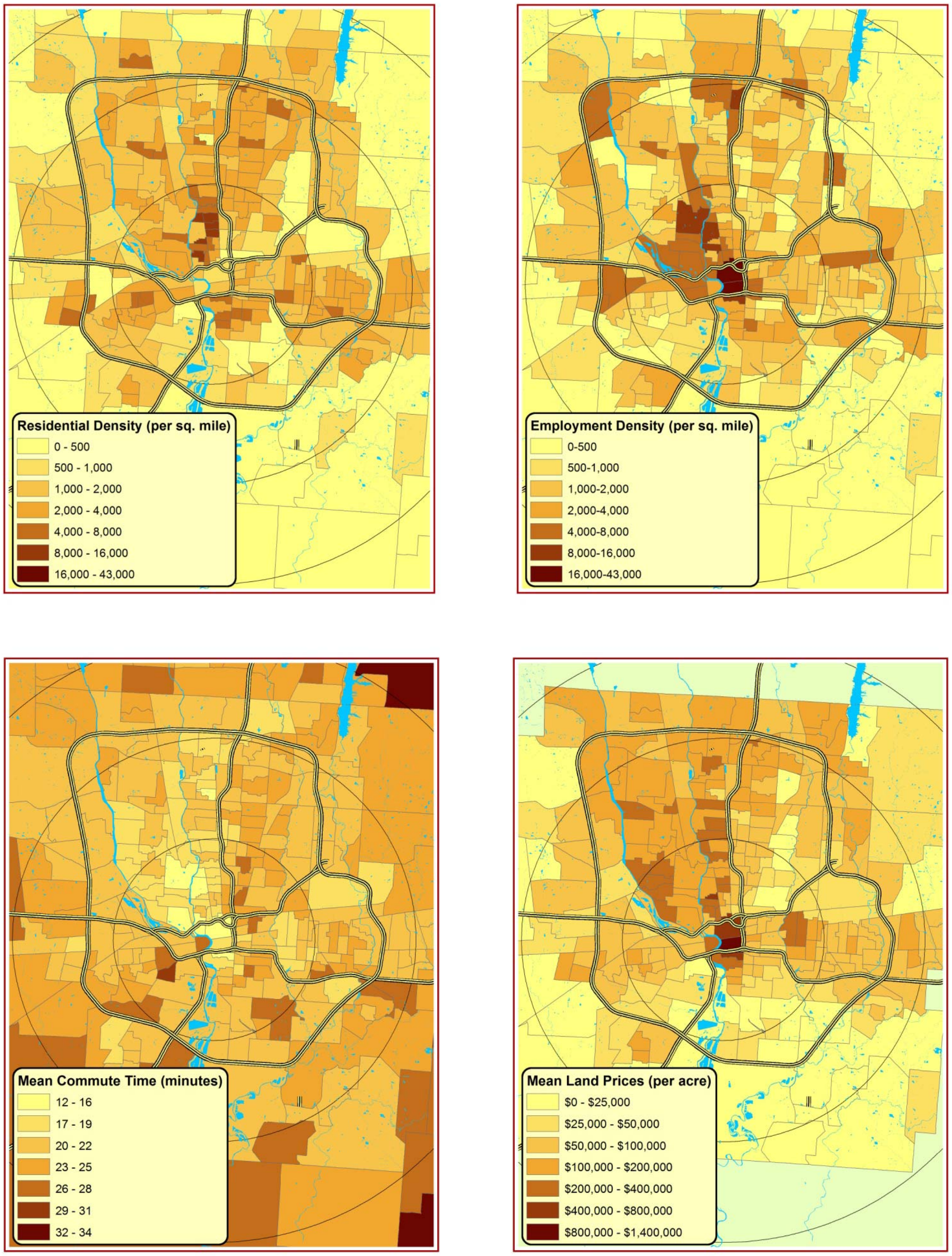

Figure 4: Spatial Characteristics of Columbus, Ohio. Rings represent five-mile increments from the city center. Densities and commute times are from the 2000 Census Transportation Planning Package. Land values are from the Franklin County Ohio Auditor's Office, 2000.

city.[27] This includes 260 census tracts, which are the unit of observation. The moments used in estimation are the sample means of each of the variables above as well as the interaction of these variables with location, r, and with net commuting, $M(r)$, which is calculated as the net jobs minus workers for the area contained in a radius, r. Overall, there are 25 moment conditions used to identify 12 parameters.[28] All observations are weighted appropriately.[29]

6 Quantitative Results

6.1 Structural Parameter Estimates

Select parameter estimates are shown in Table 2, along with 95 percent confidence intervals.[30] The consumption share parameter, which can be interpreted as the share of all consumer expenditure on all goods except land, is broadly consistent with previous estimates. Note that the land share $(1-\beta)$ is different from the housing share. Land share is the component of housing expenditures related to land as opposed to structures or improvements. This share varies over time and across cities, and it is generally lower in low-cost cities.[31] Likewise, the estimate of land share of production, $(1-\alpha)$, equal to 1.5 percent, is very much in line with the literature.[32]

As mentioned, the two parameters describing the agglomeration externality cannot be

[27]Computationally, the model is calculated out to 30 miles, well beyond the urbanized area. This is where the market-clearing condition, $M(S) = 0$, is applied. In the data, there are still 56,000 unhoused workers at this radius, which is strictly applied in the estimation.

[28]See Appendix B for details of moment conditions.

[29]For most moment conditions, the population weights are either tract-level employment or tract-level worker residential population. The alternative would be to weight by land area. However, I choose to adopt the notion supported by Glaeser and Kahn (2004), who argue that the important factor in analyzing densities and other location characteristics is the market conditions faced by individual people or firms. The exceptions are the moment conditions for land use for open space rents, where observations are weighted by land area.

[30]Certain parameters are mainly for scaling and have little practical interpretation, so they are not included here. See Appendix D for the estimates of the complete parameter vector.

[31]See Davis and Heathcote (2007) and Davis and Palumbo (2008), and Piazzesi and Tuzel (2007) for details on housing consumption.

[32]See Ciccone (2002) and Rappaport (2008) for details.

Parameter	Description	Estimate	Standard Error	95 % confidence
β	consumption share	0.979	1.11e-3	[0.977, 0.981]
α	production technology	0.985	5.58e-4	[0.984, 0.986]
$\gamma * \delta$	agglomeration externality	2.24e-3	4.27e-4	[1.39e-3, 3.09e-3]
τ	distance cost	1.06e-4	2.35e-4	[-3.64e-4, 5.77e-4]
κ	congestion cost	9.51e-9	1.89e-9	[5.74e-9, 1.33e-8]
η_C	commercial land share	0.420	1.63e-2	[0.388, 0.453]
η_R	residential land share	0.550	1.58e-2	[0.519, 0.582]
ρ	land use elasticity	6.13	1.39	[3.35, 8.91]

Table 2: Estimation results, n=260.

separately identified. The estimate of the product of the two is reported in the table. The interpretation of this parameter is approximately the percentage decline in productivity per mile. In other words, the estimate of 2.23e-3 suggests that approximately 2 percent of production is lost at a radius of 10 miles relative to the center of the city. This is hard to directly compare to the literature, given that this is a spatial measure rather than a simple elasticity. However, it is broadly consistent with numerous estimates of returns to density.[33]

A very interesting result arises in the commuting cost parameters. Notice that the distance cost parameter, τ, which can be interpreted approximately as the percentage of lost wages per mile due to commuting, is not statistically significant. The estimates suggest that the distance traveled is not a particularly important component in the cost of commuting. The point estimate of τ says that 0.1 percent of income is lost over an uncongested 10-mile commute. The congestion cost parameter, κ, on the other hand, is statistically significant and much more economically important than the distance parameter. For the level of congestion observed in these particular data, the congestion parameter estimate implies a cost of 0.87 percent of income traveling from a radius of 10 miles to the center of the city. This points to the fact that congestion, not distance, is the most important consideration in commuting decisions.

Finally, the land use parameters provide insight into the relative levels of land used for different purposes under different conditions. The two land share parameters, η_R and η_C,

[33]See Rosenthal and Strange (2004) for a summary of estimates from the literature

can be interpreted similarly to share parameters in a CES function. The estimates indicate, all things equal, that a higher share of land is dedicated to residential use. But perhaps the more interesting parameter is the elasticity parameter, ρ, which is considerably greater than unity, meaning that land use is very sensitive to relative bid rents for commercial and residential purposes. In analysis of urban spatial structure, the extensive margin matters at least as much as the intensive margin in land use.

6.2 Model Fit

The model is quite successful in fitting the underlying observed data. Figure 5 shows the model predictions plotted along side observations for key quantities and prices, as functions of radius from the center of the city. The data are averaged across 0.5-mile increments to smooth the plot for illustrative purposes. Land prices, land use intensity, and commute times are more or less log-linear in appearance with respect to radius, and the model is able to capture those gradients quite precisely. The exception seems to be in the central business district, which exhibits considerably more intensity of land use than the model predicts.

However, the real achievement of the model is that it is able to capture the rather complicated and non-monotonic land use functions, which are an important part of urban structure. The residential land use increases initially and then decreases. The commercial land use, on the other hand, decreases rapidly initially, levels out, and then decreases again.

These functions help illuminate the inner workings of the model. The land use functions are determined by the relative commercial and residential rents, which are ultimately driven by the wage gradient. Both rent functions are strictly decreasing, but the commercial rent is declining at a faster rate due to faster declines in productivity relative to wages near the center. Then wages begin to decline faster (due to higher levels of commuters), which causes the two functions to decline at a similar rate. Finally, the rents begin to approach the level of agricultural rents, which causes both commercial and residential use to fade away before agricultural use ultimately dominates. Note that these complex land use functions would not be possible if the commuting costs did not depend on congestion. If, for example, the

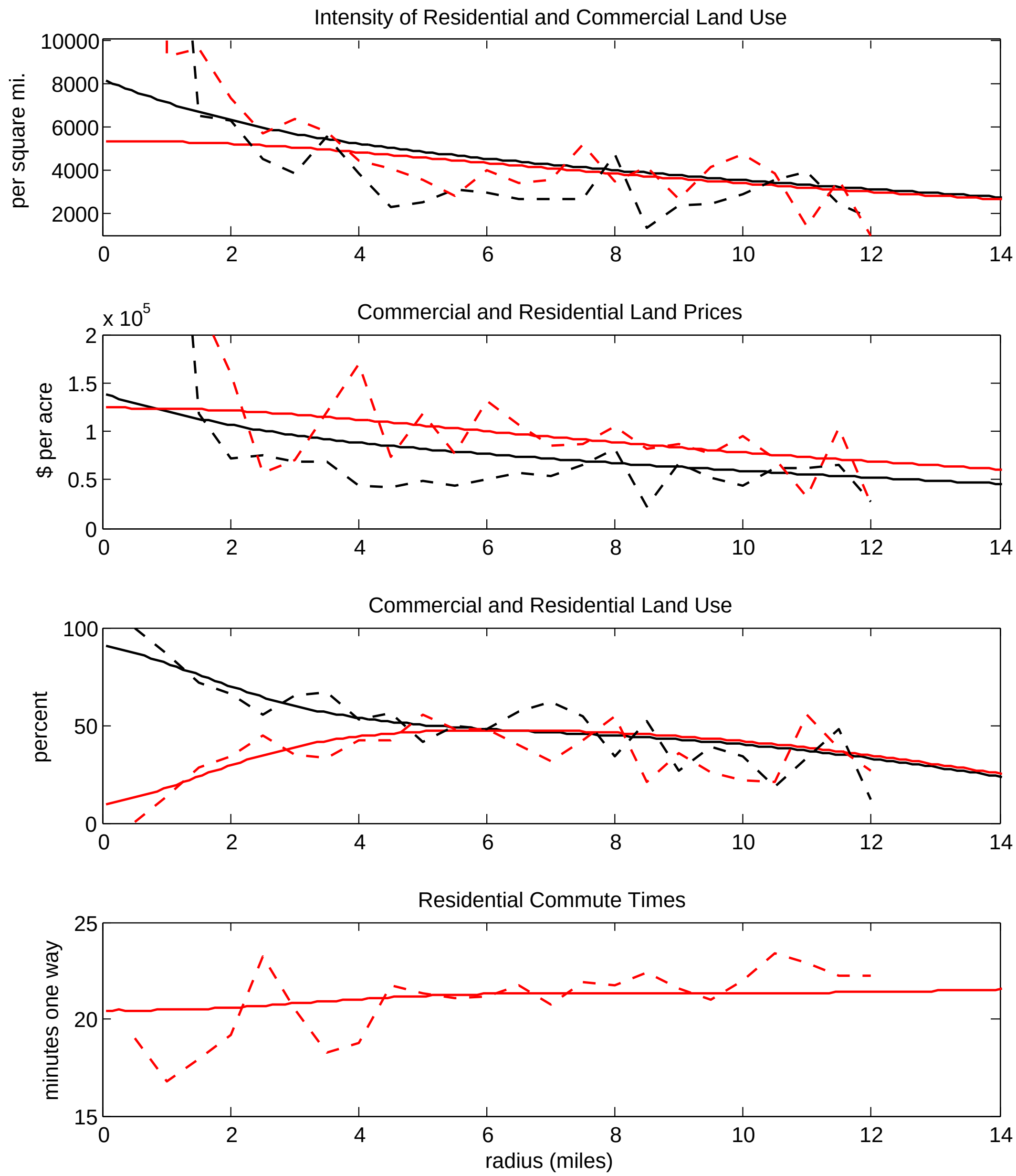

Figure 5: Model Fit: Functions calculated via the estimated computational model (solid lines) are plotted along with observations from the data averaged at 0.5-mile increments (dotted lines). The dark (black) lines represent commercial variables, while the light (red) lines represent residential variables. Data sources: Bureau of Transportation Statistics and the Franklin County Auditor's Office.

wage path declined at a constant exponential weight, the relative extent of commercial and residential land use would decline at a constant rate. The observed patterns arise because the wage path is steeper in areas with high congestion relative to areas with low congestion.

7 Comparative Statics

This section tests the effects of changes in key structural parameters of the model on outcomes in the economy. This exercise accomplishes two things. First, it tests the sensitivity of outcomes to parameter estimates. Second, it helps in understanding how different parameters affect the aggregate and spatial outcomes of the economy, in turn providing insight into the important determinants of urban spatial structure. Given that the focus of this paper is on congestion and agglomeration, the tests will include analysis of the agglomeration parameter, δ, as well as the transportation cost parameters, τ and κ, and their interaction with one another. In the experiments that follow, parameters of the model are changed and new equilibrium outcomes are calculated.

7.1 Agglomeration Attenuation

Consider the effect of changing the agglomeration attenuation parameter, δ, which determines the rate of decay of productivity spillovers.[34] The aggregate results are shown in Table 3. Given that decreasing the attenuation of agglomeration spillovers represents a technological improvement, in general the effect of decreasing δ is to increase economic activity in the aggregate. For example, total production, total employment, total wages, and total rents all increase as delta declines. This is not surprising, as the effects of production externalities are available across a wider area, given that the attenuation is lower.

The gains, however, are due mostly to increases in the size of the labor force and increases in the use of and access to land, as opposed to more efficient use of resources.

[34]Given that δ was not estimated separately from γ, for all comparitave statics, γ is set to 0.01. This ensures that $\gamma + \alpha < 1$, which is a standard stability condition. This choice can affect comparative statics quantitatively, but the qualitative analysis holds.

δ	total production ($)	total employment	per capita production ($)	production per square mi.	open space (%)
0.213	5.03E+10	1.41E+06	3.578E+04	1.30E+08	74.7
0.224	**4.18E+10**	**1.17E+06**	**3.577E+04**	**1.29E+08**	**78.7**
0.235	3.52E+10	9.83E+05	3.576E+04	1.28E+08	81.8

Table 3: Aggregate changes in the economy for changes in the agglomeration attenuation parameter, δ. Bold type represents the baseline model.

In other words, per capita production, wages, and rents all decline only slightly as delta increases. This concept is also reflected in the loss of agricultural land or open space. In the strict interpretation of this particular open city model, a decrease in delta is unambiguously welfare improving. However, if we think of a closed city where labor supply is fixed, or if we include some common-good value of open space land, the results have potential to change.

The spatial distribution of activity is not quite so straightforward. For the most part, a declining δ leads to point-wise increases at every location in the city for most of the variables of interest, including land use intensity, bid rents, wages, and also in residential land use. Commercial land use, however, decreases in the center of the city and increases on the outskirts of the city, effectively spreading production more evenly throughout the city, given the less steep gradient of the production externality, z. This increases the mixing of uses and makes residential use relatively more important in the center of the city. Figure 6 shows the spatial change in commercial land use as δ decreases.

7.2 Transportation Costs

It is also useful to revisit the effect of transportation costs. For now we will simply work with the congestion cost parameter, κ. When they are adjusted individually, the effect of the distance cost, τ, and the congestion cost, κ, are very similar. We will discuss the interaction and interpretation of these two parameters together in the next section.

When the transportation cost is decreased, the aggregate results, shown in Table 4, are fairly intuitive. First, decreasing transportation cost is a technological advancement in the economy, so all aggregate measures of economic activity increase with decreasing

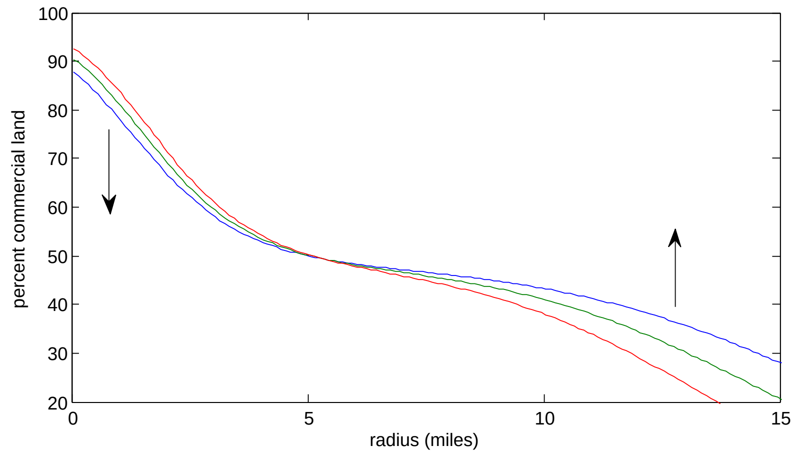

Figure 6: Changes in the percentage of land used for commercial purposes given incermental changes in agglomeration attenuation δ. Arrows represent changes in the direction of decreasing δ.

κ. This includes total production, employment, rents, and wages. Second, measures of efficient use of resources increase too, including per capita production, wages, and rent. The productivity of land increases as well. However, given that lower transportation costs result in more sparse residential development, the amount of open space decreases in the economy.

κ	total production ($)	total employment	per capita production ($)	production per square mi.	open space (%)
7.61E-09	4.35E+10	1.22E+06	3.578E+04	1.31E+08	77.9
9.51E-09	**4.18E+10**	**1.17E+06**	**3.577E+04**	**1.29E+08**	**78.7**
11.4E-09	4.06E+10	1.14E+06	3.576E+04	1.27E+08	79.2

Table 4: Aggregate changes in the economy for given incremental changes in the congestion cost parameter, κ. Bold type represents the baseline model.

When we consider the spatial distribution of activity, several patterns emerge. The most striking effect is the change in residential land use, shown in Figure 7. While residential rents and intensity increase point-wise, residential land use decreases in the center of the city and increases in the outskirts, meaning that residential distribution becomes relatively more sparse with decreasing transportation costs. Commercial land use, on the other hand, is less responsive than residential use, but the pattern is the opposite. Commercial use becomes relatively lower on the edge of the city, representing more compact distribution of jobs. This is a point that is often misunderstood. The conventional wisdom is that lower transportation costs decrease density, but this is true only with residential density; employment density has the opposite response to lower transportation costs.

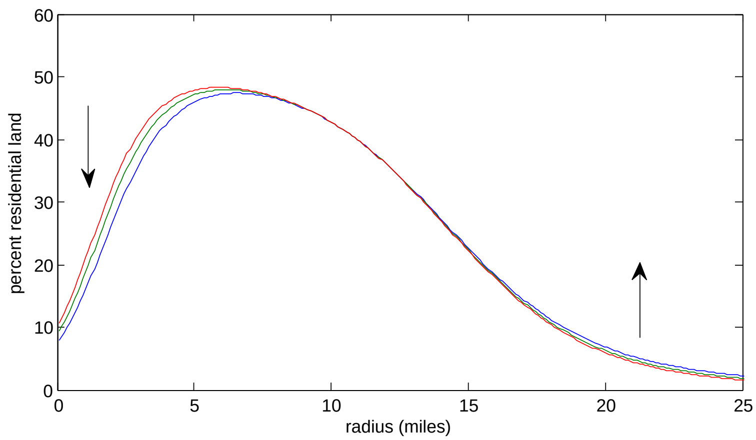

Figure 7: Changes in the percentage of land used for residential purposes for given incremental changes in congestion cost parameter κ. Arrows represent changes in the direction of decreasing κ.

When we compare the effects of changing δ versus κ, the level effects are similar, i.e., increasing the costs decreases economic measures. However, the spatial effects go in the opposite direction, for the two parameters. Increasing the attenuation of agglomeration concentrates employment and disperses residents, where as increasing transportation costs disperses employment and concentrates residents. These are intuitive results, and not necessarily novel, but it is useful to confirm and rehash these ideas within the current general equilibrium framework.

7.3 Transportation Technologies

Next, we consider how the two transportation parameters, τ and κ, affect the economy differently. Recall that τ is a measure of the cost of commuting over an uncongested distance, while κ captures the part of the commuting cost that depends on the level of congestion.

In some sense, these parameters can be interpreted as transportation technology parameters. Using two different parameters allows us to think about transportation provision on two dimensions. In other words, these parameters represent the efficiency of different transportation infrastructure or planning under different congestion conditions. For example, underground subways are very efficient in congested areas, but less so in exurban or sparse areas. Automobiles on two-lane roads would probably have the opposite characteristic, being very efficient in rural areas but less so in city centers. Likewise, the distribution of transportation provision in different parts of a city would change the cost equation in different ways. Many other factors, from parking regulations to network connectivity, will change this equation.

Nonetheless, for simplicity, from here forward, we will associate higher congestion costs and lower distance costs with automobile infrastructure, and lower congestion costs and higher distance costs with public transit infrastructure. This is clearly an extreme oversimplification of transportation network design, but it allows us to think about possible policy implications related to transportation provision in a tangible and accessible way, within the context of a multi-agent open city model.

One way to examine the influence of transportation technologies on spatial structure is to hold all other parameters constant, adjust the transportation technologies, and analyze their effects. The goal of these experiments is not to analyze an increase or decrease in total transportation costs, but instead to understand how the mix of transportation technologies changes the structure of the city. We cannot truly study the welfare ramifications, given that we do not know the costs of different transportation provision. However, if we assume that the policy experiments are cost neutral, then, strictly speaking, the total land rent in the city is a measure of welfare, the intuition being that the value of the fixed resource is maximized.[35] We will also consider other things like productivity and land use.

In order to maintain consistency and anchor the analysis, each counterfactual experi-

[35] See Rossi-Hansberg (2004) for a detailed discussion of welfare and optimal policy in a similar setting.

ment is employment neutral, meaning that total employment in the city remains unchanged. In order to maintain employment levels, if we raise the distance cost, we must lower the congestion cost, corresponding to an increase in transit spending at the expense of highways. Furthermore, we will consider the initial estimates to be a baseline model from a purely automobile-oriented city, given the low estimate for distance costs and the fact that the model was estimated using a city with very little transit. Policies demonstrated here incrementally increase distance costs and reduce congestion costs to illustrate the effect of gradually increasing transit provision.

The aggregate results, in Table 5 show positive consequences of trading higher distance costs for lower congestion costs. Production and rents both increase, while more land is preserved for open space. In addition, the economy exhibits both higher production and increased efficiency in the use of resources, in terms of production per unit of land and per worker. The more striking result is that these gains are made despite the fact that total congestion and total commuting costs increase. The intuition is that workers are more willing to commute into congested areas given the lower costs; therefore firms can cluster and take advantage of the production increases from proximity. These increases in productivity offset the increases in transportation costs. The upshot is that congestion is not always a bad thing. It can be, in fact, evidence of productivity.

κ	τ	per capita production ($)	total rent ($)	open space (%)	total commuting (person miles)
9.51E-09	**1.06E-04**	**3.5769E+04**	**1.70095E+09**	**75.79**	**2.81E+06**
4.76E-09	6.05E-04	3.5773E+04	1.70102E+09	75.83	2.95E+06
0	1.21E-03	3.5782E+04	1.70136E+09	75.91	3.26E+06

Table 5: Aggregate changes in the economy for given incremental changes in the transportation technology parameters, κ and τ, holding total employment constant. Bold type represents the baseline model.

The spatial effects of changing the transportation technology are a little harder to characterize. The changes in all variables of interest are decidedly non-monotonic with respect to the distance from the center of the city, and the results are fairly nuanced. However, the intuition described above, that businesses become more clustered, is the most important takeaway. This is perhaps best illustrated by the changes in total employment at each location, shown in Figure 8. Employment decreases somewhat in the very center of the city, but increases between 3 and 10 miles from the center and falls off in the outskirts before fading to zero. The net effect is that production is, on average, more centrally located, boosting production externalities, and productivity increases everywhere.

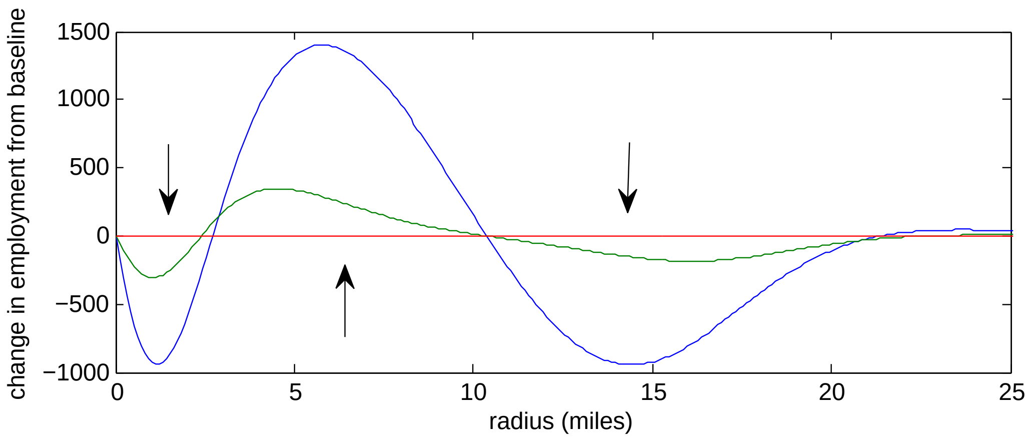

Figure 8: Changes in total employment with changes in transportation technology. Arrows represent changes in the direction of decreasing the congestion cost κ and increasing the distance cost τ. Total employment is held constant.

8 Congestion Pricing: A Policy Simulation

We could explore any number of policy experiments within the framework presented here, including, but not limited to, development subsidies, tax codes, transportation provision, or zoning. However, this model is particularly well suited to study the effects of congestion pricing on urban structure and economic outcomes. In this section, we explore the results of congestion pricing by implementing a naive optimal congestion tax and solving for equilibrium outcomes computationally, within the framework of our model.

The conventional wisdom is that congestion pricing can improve welfare by internalizing the external costs associated with increased congestion. As mentioned, this is not just a theoretical exercise; congestion pricing has already been implemented in cities and is often discussed as a solution to increasing traffic congestion and declining revenues from gas taxes.

The theory of congestion pricing is simple and can be modeled as a standard Pigovian tax. The goal is to implement a tax that is equal to the external social cost of an activity. In this case, the marginal social cost of an additional commuter is simply the congestion cost parameter, κ, given that the cost function is linear in the level of commuters, M(r). The total congestion cost applies to all other commuters at location r, so the external cost and hence the optimal congestion tax through a location r is

$$Congestion\ Tax = \kappa\, M(r)\, w(r).$$

In other words, the optimal tax doubles the congestion costs faced by a commuter. The harder problem, in both theory and implementation, is how to redistribute the revenue. In order to avoid distortion, in this case, we can simply provide a lump-sum subsidy to wages that is constant at every location r. The amount of revenue available for redistribution is dependent on the total congestion, which simultaneously depends on the wage subsidy, so there is no analytic solution of which I am aware. However, it is straightforward to solve for the total tax revenue computationally, using either a search algorithm or a fixed-point iteration.

Of course, given that there are also agglomeration externalities, it becomes clear that congestion pricing will not be as effective as expected, given that the policy will reduce concentration of employment and therefore productivity. In fact, the policy could even have net negative effects if the loss of productivity completely offsets the gains from reduced congestion.

I compare the effects of optimal congestion pricing for two separate experiments relative to the baseline estimated model. The first is denoted as the "no agglomeration" effect of congestion pricing. In this case, I consider what the effect of congestion pricing would be if location-specific productivity, $z(r)$, were fixed and not dependent on the spatial distribution of employment. In some sense, this is the conventional view of how congestion pricing is supposed to work and ignores the positive benefits of agglomeration externalities. To implement this policy, I simply keep the function $z(r)$ constant while calculating the congestion price and wage subsidy. Calculating these results, provides a sense of how different the estimates of the benefits of congestion prices can be if agglomeration externalities are ignored. The second experiment, denoted "agglomeration," includes all of the equilibrium effects of congestion pricing, including the loss of productivity resulting from employment dispersion. This represents the true effect of congestion pricing when agglomeration externalities are included.

The aggregate results of the two different policies are shown in Table 6, along with the baseline comparison. First, let us compare the "no agglomeration" congestion pricing to the baseline model. In this case, it appears that big gains can be made by implementing congestion pricing in terms of total production, total employment, and total land rent. Most of the gains come from increased employment and land use as opposed to higher factor productivity. Both total congestion and total commuting costs decline substantially, which is not surprising given that congestion is the target of the tax. Open space declines in this scenario, due to increased population and dispersion. Overall, these results confirm the theory that taxes on negative externalities can improve outcomes in an economy.

	total production ($)	total employment	total rent ($)	open space (%)	total commuting (person miles)
Baseline	4.18E+10	1.1694E+06	1.703E+09	0.759	2.817E+06
no agglomeration	4.39E+10	1.2288E+06	1.776E+09	0.750	1.719E+06
agglomeration	4.17E+10	1.1687E+06	1.702E+09	0.760	1.716E+06

Table 6: Aggregate effects of naive optimal congestion pricing for the case where location-specific productivity is assumed fixed (no agglomeration) and the full-equilibrium case where location productivity is endogenous and depends on employment, (agglomeration).

Now let us consider the full equilibrium effect with agglomeration externalities. The striking feature of the results is that there is only a small effect on important economic indicators when the congestion pricing is implemented, and in fact there are slight declines in production, employment, and total rent, meaning that there is a net negative effect of congestion pricing on welfare. There is still a strong effect on congestion and commuting costs, which is not surprising, but these gains are totally offset by the loss in productivity due to the dispersion of employment. This is illustrated in Figure 9, which shows the change in employment by location for the two policy simulations relative to the baseline. Employment declines in the center of the city and increases at its edge. Note that employment in the "no agglomeration" case is higher everywhere compared to the case where endogenous agglomeration is included, given that there is no loss of productivity due to dispersion in the former case.

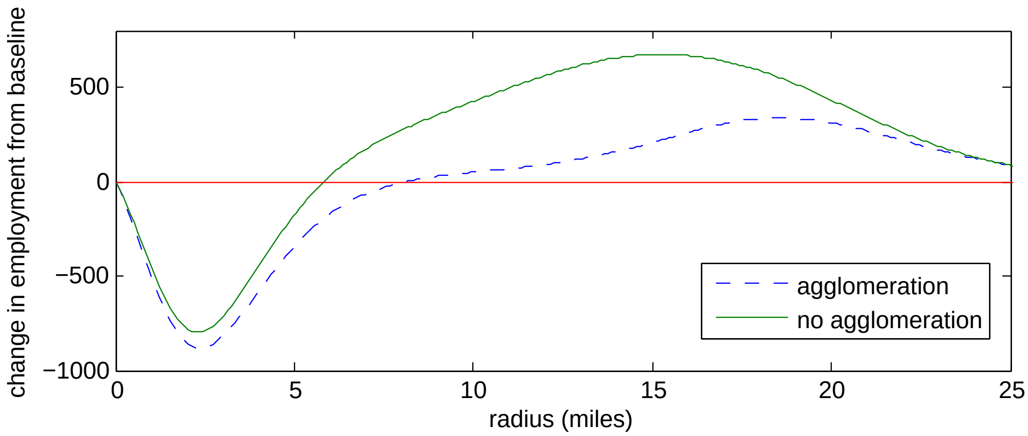

Figure 9: Spatial changes in employment for naive optimal congestion pricing for the case where location-specific productivity is assumed fixed (no agglomeration) and the full equilibrium case where location productivity is endogenous and depends on employment (agglomeration).

It is important to note that these results are very sensitive to the parameter estimates, particularly the relative strengths of the agglomeration attenuation parameter, δ, and the transportation cost parameter, κ. For parameter values well within the confidence intervals of the estimates, the net effect on economic variables can be either positive or negative. However, the net gains from congestion prices are muted significantly for any reasonable range of parameters as compared to the "no agglomeration" outcome.

Several additional caveats are in order, which lend caution to completely eradicating congestion pricing from the list of viable policy alternatives. First, this analysis includes only time costs of congestion, and there are other costs, to health and environment in particular, related to congestion. Second, the lost productivity due to congestion pricing could be mitigated if commuters have flexibility on other dimensions of commuting behavior. For example, commuters may be able to switch modes, which would offset some of the costs associated with congestion pricing and therefore reduce the response in the location choice of firms and workers. However, given that public transit is not available for large sections of the population, particularly in the U.S., it is unclear that this would have a strong mitigating effect. Another possibility is that commuters could change the time of day that they travel to work. Again, it is unclear what costs are associated with changing departure times, and it is entirely likely that those commuters most sensitive to pricing are the same commuters that have the least amount of flexibility in their work schedules.

None of these additional considerations contradicts the main conclusion of the analysis, which is that the benefits of congestion pricing schemes will be muted in the presence of agglomeration externalities. Furthermore, under certain conditions, congestion pricing could even lead to net negative consequences for welfare. Policymakers should be aware of this fact.

9 Conclusions

In this paper, I have presented and analyzed a model of urban structure that includes both a negative externality in the form of congestion costs and a positive externality in the form of agglomeration spillovers. First, I established some basic facts about the observed structure of urban areas. Some of these features are familiar, such as the fact that employment is relatively more clustered than population, and that wages are correlated with commute times. However, the analysis also illustrated and explained land use patterns in urban areas, including the extent to which uses are mixed and the observed gradual transitions between uses. These land use characteristics are clearly important features of urban structure, but are often not well-explained by urban spatial models.

By estimating the structural parameters of the model and matching the observed characteristics of an urban economy, I am able to characterize the relative magnitudes of two related but offsetting externalities. This allowed for the calculation of counterfactuals, in a full equilibrium set-up, that showed the effect of transportation costs and agglomeration externalities on the spatial distribution of economic activity in urban areas.

The most striking result of the analysis is that, using estimated parameters, policy simulations showed that a perceived optimal congestion pricing policy can lead to net negative economic outcomes. Admittedly, further investigation is needed to fully vet these results. However, these outcomes illustrate and provide evidence that the unintended consequences of congestion pricing, or any other urban policy targeted toward urban mobility or land use, should be in the minds of policymakers, given the strong evidence for congestion costs and agglomeration externalities. More generally, these results suggest that policy analysis can be complex in other economies characterized by multiple externalities.

This analysis leaves room for many extensions. Certainly, more complexity can be added. In particular, agent heterogeneity is an important consideration given that we know that firms and employees sort spatially in urban areas. This has consequences for estimation and for policy analysis. More interesting, perhaps, is that the current analysis has potential to help better understand the dynamics of urban structure. Cities have undergone substantial changes in terms of their spatial structure over time. The trend of suburbanization in the last century is well known, and it has been linked to declining transportation costs. However, evidence is increasing that, over the last decade or so, there has been a growing trend toward reurbanization of residential populations in U.S. cities, with central cities showing growth for the first time in over a half a century. Furthermore, little is known about the changing spatial allocation of employment, or how structural changes in sectoral composition across or within urban areas are related to urban structure. The framework presented here could be extended in a straightforward way to answer these questions.

References

Ahlfedlt, G., Redding, S. J., Sturm, D., & Wolf, N. (2013). The economics of density: Evidence from the Berlin wall. *Econometrica*.

Anas, A., Arnott, R., & Small, K. A. (1998). Urban spatial structure. *Journal of Economic Literature*, *36*.

Anas, A., & Kim, I. (1996). General equilibrium models of polycentric urban land use with endogenous congestion and job agglomeration. *Journal of Urban Economics*, *40*(2).

Arnott, R. (2007). Congestion tolling with agglomeration externalities. *Journal of Urban Economics*, *62*.

Arnott, R., DePalma, A., & Lindsey, R. (1993). A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *American Economic Review*, *83*.

Arzaghi, M., & Henderson, J. V. (2008). Networking off Madison Avenue. *The Review of Economic Studies*, *75*.

Baum-Snow, N. (2007). Did highways cause suburbanization?. *Quarterly Journal of Economics*, *122*.

Brinkman, J. C., Coen-Pirani, D., & Sieg, H. (2012). Firm dynamics in an urban economy. *Federal Reserve Bank of Philadelphia Working Paper*.

Brueckner, J. K. (2002). Airport congestion when carriers have market power. *American Economic Review*, *92*.

Burchfield, M., Overman, H. G., Puga, D., & Turner, M. A. (2006). Causes of sprawl: A portrait from space. *Quarterly Journal of Economics*.

Chatterjee, S., & Eyigungor, B. (2012). A tractable circular city model with an application to the effects of development constraints on land rents. *Federal Reserve Bank of Philadelphia Working Paper*.

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

Davis, M. A., Fisher, J. D., & Whited, T. M. (2011). Macroeconomic implications of agglomeration. *SSRN Working Paper*.

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

Davis, M. A., & Palumbo, M. G. (2008). The price of residential land in large US cities. *Journal of Urban Economics*, *63*.

Duranton, G., & Puga, D. (2001). Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review*, *91*(5).

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. *Handbook of Regional and Urban Economics*, *4*.

Duranton, G., & Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities. *American Economic Review*, *101*.

Ellison, G., & Glaeser, E. L. (1997). Geographic concentration in us manufacturing industries: A dartboard approach. *Journal of Political Economy*, *105*.

Fujita, M., & Ogawa, H. (1983). Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional science and urban economics*, *12*.

Fujita, M., & Thisse, J.-F. (2002). *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*. Cambridge.

Glaeser, E. L., & Kahn, M. E. (2004). Sprawl and urban growth. *Handbook of Regional and Urban Economics*, *4*.

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

Henderson, V., Kunkoro, A., & Turner, M. (1995). Industrial development in cities. *Journal of Political Economy*, *103*(5).

Holmes, T. J. (2005). The location of sales offices and the attraction of cities. *Journal of Political Economy*, *113*.

Jacobs, J. (1969). *The Economy of Cities*. Random House.

Lucas, R. E., & Rossi-Hansberg, E. (2002). On the internal structure of cities. *Econometrica*, *70*(4).

Marshall, A. (1890). *Principles of Economics*. Macmillian.

Mayer, C., & Sinai, T. (2003). Network effects, congestion externalities, and air traffic delays: Or why not all delays are evil. *American Economic Review*, *93*.

Mills, E. S. (1967). Transportation and patterns of urban development: An aggregative model of resource allocation in a metropolitan area. *American Economic Review*, *57*(2).

Ng, C. F. (2012). Heterogeneous households and firms in an urban model with open space and agglomeration economies. *Papers in Regional Science*.

Piazzesi, Monika, M. S., & Tuzel, S. (2007). Housing, consumption and asset pricing. *Journal of Financial Economics*, *83*.

Rappaport, J. (2008). Consumption amenities and city population density. *Regional Science and Urban Economics*, *38*.

Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *The Journal of Urban Economics*, *50*(2).

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

Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. *Handbook of Regional and Urban Economics*, *4*.

Rossi-Hansberg, E. (2004). Optimal urban land use and zoning. *Review of Economic Dynamics*, *7*.

Small, K. A. (1983). The incidence of congestion tolls on urban highways. *Journal of Urban Economics*, *13*.

Small, K. A., & Verhoef, E. T. (2007). *The Economics of Urban Transportation*. Routledge.

Timothy, D., & Wheaton, W. C. (2001). Intra-urban wage variation, employment location, and commuting times. *Journal of Urban Economics*, *50*.

Tse, C. Y., & Chan, A. W. (2003). Estimating the commuting cost and commuting time property price gradients. *Regional Science and Urban Economics*, *33*.

Von Thünen, J. H. (1826). *The Isolated State*. Oxford.

Wheaton, W. C. (2004). Commuting, congestion, and employment dispersal, in cities with mixed land use. *Journal of Urban Economics*, *55*.

A Existence and Uniqueness

Start with the idea that the economy is uniquely characterized by two functions defined on r: the productivity function $z(r)$ and the wage path $w(r)$. This should be clear given that all other outcomes, including wages, rents, densities, and land use, are functions of these two variables if, in fact, agents solve their maximization problem and land markets clear. Proving uniqueness then reduces to the problem of showing that for a given parameter vector there is a unique pair $\{z(r), w(r)\}$, such that the labor market clears, $M(S) = 0$, the wage path restriction, equation (7), holds, and the productivity function is internally consistent in that it satisfies equation (1). To start, I will assume $z(r)$ is fixed and show that there is only one wage path $w(r)$ consistent with the labor-market-clearing condition, M(S)=0.

Proposition 1 *Given that $N_C(w(r), z(r)) = n_C(w(r), z(r))\theta_C(w(r), z(r))$ is strictly decreasing in $w(r)$, $N_R(w(r), z(r)) = n_R(w(r))\theta_R(w(r), z(r))$ is strictly increasing in $w(r)$, and $T'(M(w(r), z(r)))$ is increasing in $M(r)$. For a fixed $z(r)$, such that $0 < z(r) < \infty$,* *there exists a unique wage path, $w(r)$, such that that the labor market clearing condition, $M(S) = 0$, holds.*

Proof: First, note that given the assumptions on N_C and N_R, then $\partial M(w(r), z(r))/\partial r = N_C(w(r), z(r)) - N_R(w(r), z(r))$ is decreasing in $w(r)$. That is, the change in the number of commuters at a location, r, is decreasing in wages at a location r. Consider an equilibrium wage path $w^*(r)$ such that $M(S) = 0$. At the center of the city, $w^*(0)$ will produce a given $M^*(0)$. By assumption, if $w(0) < w^*(0)$, then $M(0) > M^*(0)$. Given that $T'(M(w(r), z(r)))$ is increasing in $M(r)$ and given the wage path restriction from equation (7), a higher $M(0)$ will lead to a lower $w(\epsilon)$ relative to $w^*(\epsilon)$, a small distance, ϵ, from the center. Therefore $M(\epsilon) > M^*(\epsilon)$. Carrying this out to the edge of the city, it is straightforward that $M(S) > M^*(S)$. We can use the same argument to show that for $w(0) > w^*(0)$, $M^*(S) > M(S)$. It is also trivial that for $w(0) \to \infty$, $M(S) < 0$, and for $w(0) \to -\infty$, $M(S) > 0$, given that $z(r)$ is positive and finite. Therefore, for any fixed productivity function, $0 < z(r) < \infty$, there exists a unique wage path, $w^*(r)$, such that $M(S) = 0$.

QED

Next we need to show that, for a given parameter vector, there is a unique $z(r)$, such that equation (1) holds. Note that the specification for the externality $z(r)$ is determined by solely by $z(0)$, and point-wise, $z(r)$ increases with $z(0)$. Also note that, that for a given $z(0)$, it is straightforward to calculate the outcomes of the entire economy consistent with all other equilibrium conditions, save the externality specification. This suggests a fixed-point problem of the following form,

$$z^*(0) = Z(z^*(0)),$$

where $z^*(0)$ is an equilibrium solution. Therefore, if there is a unique fixed point for the above equation for a given parameter set, then we also have a unique equilibrium. In standard agglomeration models, using Cobb-Douglas forms, a sufficient condition for uniqueness is $\alpha + \gamma < 1$, which ensures that the mapping of z onto itself is concave. This condition holds in the estimated parameter vector. We cannot be sure that this is sufficient in the current set-up given the endogenous congestion function and nonlinear land use function. However, given that the fixed point operator is dependent only on $z(0)$, it is quite simple to computationally plot the mapping of $z(0)$ onto itself to confirm that the function is well-behaved. This result is shown in Figure 10, with the equilibrium $z^*(0)$ normalized to 1. This provides strong evidence that the mapping is well-behaved and that the equilibrium is globally unique.

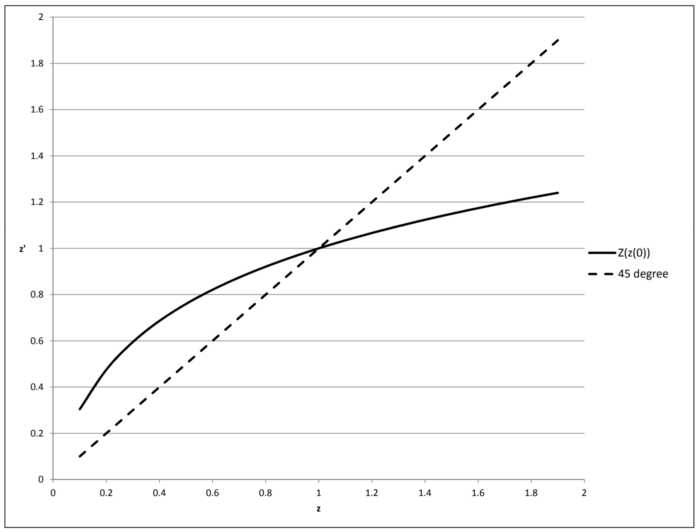

Figure 10: Illustration of fixed-point operator $Z(z(0))$, calculated computationally for the estimated parameter vector

B Moment Conditions

Moment conditions used in the estimation are shown below. Predicted values from the computational model are denoted with a hat ($\hat{}$) to distinguish them from sample equivalents. $T_R(r)$ represents the commuting cost by residential location. There are 12 parameters and 25 overidentifying restrictions:

1. $E[\ln(n_R(r)) - \ln(\widehat{n}_R(r;\Theta))] = 0$

2. $E[(\ln(n_R(r)) - \ln(\widehat{n}_R(r;\Theta)))r] = 0$

3. $E[(\ln(n_R(r)) - \ln(\widehat{n}_R(r;\Theta)))M(r)\ r] = 0$

4. $E[\ln(q_R(r)) - \ln(\widehat{q}_R(r;\Theta))] = 0$

5. $E[(\ln(q_R(r)) - \ln(\widehat{q}_R(r;\Theta)))r] = 0$

6. $E[(\ln(q_R(r)) - \ln(\widehat{q}_R(r;\Theta)))M(r)\ r] = 0$

7. $E[\ln(n_C(r)) - \ln(\widehat{n}_C(r;\Theta))] = 0$

8. $E[(\ln(n_C(r)) - \ln(\widehat{n}_C(r;\Theta)))r] = 0$

9. $E[(\ln(n_C(r)) - \ln(\widehat{n}_C(r;\Theta)))M(r)\ r] = 0$

10. $E[\ln(q_C(r)) - \ln(\widehat{q}_C(r;\Theta))] = 0$

11. $E[(\ln(q_C(r)) - \ln(\widehat{q}_C(r;\Theta)))r] = 0$

12. $E[(\ln(q_C(r)) - \ln(\widehat{q}_C(r;\Theta)))M(r)\ r] = 0$

13. $E[\ln(\theta_R(r)) - \ln(\widehat{\theta}_R(r;\Theta))] = 0$

14. $E[(\ln(\theta_R(r)) - \ln(\widehat{\theta}_R(r;\Theta)))r] = 0$

15. $E[(\ln(\theta_R(r)) - \ln(\widehat{\theta}_R(r;\Theta)))M(r)\ r] = 0$

16. $E[\ln(\theta_C(r)) - \ln(\widehat{\theta}_C(r;\Theta))] = 0$

17. $E[(\ln(\theta_C(r) - \ln(\widehat{\theta}_C(r;\Theta))r] = 0$

18. $E[(\ln(\theta_C(r) - \ln(\widehat{\theta}_C(r;\Theta))M(r)\ r] = 0$

19. $E[\ln(q_F) - \ln(\widehat{q}_F(\Theta))] = 0$

20. $E[\ln(T_R(r)) - \ln(\widehat{T}_R(r;\Theta))] = 0$

21. $E[(\ln(T_R(r)) - \ln(\widehat{T}_R(r;\Theta)))r] = 0$

22. $E[(\ln(T_R(r)) - \ln(\widehat{T}_R(r;\Theta)))M(r)\ r] = 0$

23. $E[(M(r)) - (\widehat{M}(r;\Theta))] = 0$

24. $E[((M(r)) - (\widehat{M}(r;\Theta)))(n_C(r)\ \ \theta_C(r) - n_R(r)\ \ \theta_R(r))\ r] = 0$

25. $E[(w(r)) - (\widehat{w}(r;\Theta))] = 0$

C Commuting Cost by Residential Location

The equilibrium solution pins down aggregate commuting through a given location, but does not identify the residential origin of the commuters. To provide intuition on this point, denote the mass of commuters moving through a location, r, as $M(r)$. When these commuters arrive at a given location r, they are indifferent between filling a job at that location or commuting on, given that they are exactly compensated for additional commuting with additional wages. Therefore, the model does not specify which commuters fill jobs at any given location.

This means that the average commuting cost is not defined by residential location in the model. Empirically, this creates a problem given that commute times are reported by residential location. To solve this problem, we need a rule that defines how jobs are filled at any given location. Given that commuters are indifferent, an obvious assumption is that commuters fill jobs with equal probability regardless of residential origin.

Assumption 8 *At any given job location, r, commuters traveling through r are equally likely to fill available jobs, regardless of residential origin.*

This implies that the proportion of jobs filled at any given location, r', by commuters originating from r is equal to the ratio of the mass of commuters originating from r to the total mass of commuters at r'. If we denote the mass of commuters originating at r who pass through r' as $M_r(r, r')$, then the following differential equation defines the change in $M_r(r, r')$, given $M(r')$:

$$\frac{\partial M_r(r, r')}{\partial r'} = \frac{M_r(r, r')}{M(r')} \left(2\pi r' n(r')(\theta(r')\right) .$$

This differential equation, along with initial conditions at a small epsilon inside the edge of the city,

$$M(S - \epsilon) = M_r(S - \epsilon, S - \epsilon) = 2\pi\, S\ \left(N(S)(1 - \theta(S)) - n(S)\theta(s)\right) ,$$

allows for computation (numerically) of the commuting patterns for all workers by residential location. Given the commuting patterns by residential location, the average commuting cost by residential location is given by

$$\frac{1}{2\pi r N(r)(1-\theta(r))}\int_0^r M(r,r')(\tau+\kappa m(r'))dr'.$$

Finally, one additional adjustment needs to be made to the commuting data. While the model produces zero commuting costs for some locations commute times in the data are never zero. This implies that there is some fixed commuting cost that is present for all workers. In other words, if one were to regress commute times on distance from the center of the city, there is an obviously significant intercept term. This can be modeled with an additional parameter, c_f, interpreted as a fixed cost. It is straightforward that this parameter is equivalent to a shift in wages for all workers, but it is useful to apply the cost directly to commuting to match the model and data.

D Complete Parameter Estimates

Parameter	Description	Estimate	Standard Error	95 % confidence
β	consumption share	0.979	1.11e-3	[0.977, 0.981]
α	production technology	0.985	5.58e-4	[0.984, 0.986]
$\gamma * \delta$	agglomeration externality	2.24e-3	4.27e-4	[1.39e-3, 3.09e-3]
τ	distance cost	1.06e-4	2.35e-4	[-3.64e-4, 5.77e-4]
κ	congestion cost	9.51e-9	1.89e-9	[5.74e-9, 1.33e-8]
η_C	commercial land share	0.420	1.63e-2	[0.388, 0.453]
η_R	residential land share	0.550	1.58e-2	[0.519, 0.582]
ρ	land use elasticity	6.13	1.39	[3.35, 8.91]
u	reservation utility	2.33e4	5.98e2	[2.21e4, 2.45e4]
A	total factor productivity	4.13e4	4.76e2	[4.04e4, 4.23e4]
q_F	open space rent ($/sq.mi./yr)	1.18e5	1.32e4	[9.17e4, 1.45e5]
c_f	fixed commuting cost	1.50e3	1.82e1	[1.46e3, 1.54e3]

Table 7: Estimation results, n=260