# A Unified Theory of Cities*

Jacques-François Thisse, Matthew A. Turner, Philip Ushchev

November 2021

Abstract: *How do people arrange themselves when they are free to choose work and residence locations, when commuting is costly, and when increasing returns may affect production? We consider this problem when the location set is discrete and households have heterogenous preferences over workplace-residence pairs. We provide a general characterization of equilibrium throughout the parameter space. The introduction of preference heterogeneity into an otherwise conventional urban model fundamentally changes equilibrium behavior. Multiple equilibria are pervasive although stable equilibria need not exist. Stronger increasing returns to scale need not concentrate economic activity and lower commuting costs need not disperse it. The qualitative behavior of the model as returns to scale increase accords with changes in the patterns of urbanization observed in the Western world between the pre-industrial period and the present.*

JEL: Ro

---

* We gratefully acknowledge helpful conversations with Dan Bogart about the history of cities and agglomeration effects and with Kiminori Matsuyama about the concept of stability. We also thank Jonathan Dingel, Gilles Duranton, Frédéric Robert-Nicoud, Andrii Parkhomenko, Giacomo Ponzetto, Steve Redding, Roman Zarate, and participants at HSE, the Online Spatial and Urban Seminar, Princeton, and the 10th European Meeting of the Urban Economics Association for useful comments. Turner is grateful for the support of a Kenen fellowship at Princeton University during part of the time this research was conducted. Thisse and Ushchev acknowledge the support of the HSE University Basic Research Program.

Two centuries ago, one human out of ten lived in a city, incomes were a fraction of their present level, and commuting occurred on foot. Today, more than half of the world's population is urbanized, and commuting by foot is a curiosity in much of the world. Conventional wisdom holds that this process of urbanization is a contest between increasing returns to scale in production and the cost of commuting. Returns to scale in production leads to highly concentrated employment while commuting allows people to live at lower densities than those at which they work: as returns to scale increase and commuting costs fall, cities grow.

We investigate how the organization of cities is related to the strength of returns to scale and to the cost of commuting. Our problem is to understand how people arrange themselves when they are free to choose work and residence locations, when commuting is costly, and when increasing returns may affect production. We consider this problem in a framework with a stylized discrete geography, local increasing returns to scale, and households with heterogenous preferences over workplace-residence pairs. We characterize equilibrium throughout the parameter space. More specifically, we describe the implications of arbitrarily strong increasing returns to scale for the organization of cities; demonstrate highly non-monotone comparative statics for commuting costs and preference heterogeneity; find that corner equilibria are ubiquitous; and, provide an algorithm to evaluate the stability of any equilibria.

The contest between increasing returns to scale and commuting costs does not play out as the conventional wisdom suggests. Qualitative features of an equilibrium city depend sensitively on the strength of returns to scale in production. When production is not constant returns to scale, corner equilibria, where production concentrates in a few locations, must always occur. With constant returns to scale or weak increasing returns, there is a unique interior equilibrium, economic activity is centralized, and stronger returns to scale increase central employment, wages and land rents. When increasing returns to scale are moderate, multiple stable interior equilibria may occur and stronger returns to scale may decrease central employment, wages and land rents. When returns to scale are strong, stronger increasing returns to scale disperse employment and equalize wages and land rent across locations. That is, the conventional intuition that returns to scale are an agglomeration force does not hold for sufficiently strong increasing returns to scale.

Our findings about commuting costs are equally surprising. We find that employment is dispersed both when commuting costs are high and when they are low. It is only at intermediate levels of commuting costs that highly concentrated employment can arise. That is, when the location of production and residence is endogenous, the standard intuition that decreases in commuting costs must disperse economic activity need not apply.

Our investigation is important for a number of reasons. We address a foundational problem of urban economics. While much progress has been made on this problem, existing work relies on strong simplifying assumptions and arbitrary restrictions on the strength of agglomeration forces. For example, the workhorse monocentric city model sets the location of work exogenously, while recent developments based on quantitative spatial models require assumptions prohibiting multiple equilibria. Apart from its stylized geography, our model is general and allows us to provide the complete analytical characterization of equilibrium that has long eluded the literature.

Common theoretical approaches to the economics of cities can be usefully divided into two frameworks, classical urban economics and quantitative spatial modeling (QSM). The classical urban economics literature assumes homogenous households and stylized continuous geographies. In contrast, QSM assumes agents with heterogeneous preferences over workplace-residence pairs; complex, empirically founded discrete geographies; and a flexible description of the first nature advantages of particular locations for work or residence. We consider a hybrid case. We apply a quantitative spatial model to a stylized geography, a discrete linear city on a featureless landscape, much like what is often studied in the classical approach. A single parameter describes household heterogeneity in our model, as in the QSM literature, and so we are able to investigate what happens as the heterogenous households approach the homogeneity of the older urban economics literature. In this way, we unify the two literatures and allow for a more complete understanding of both classes of models and their different implications for how cities behave.

The flexibility of QSM permits an analysis of real world comparative statics in estimated or calibrated models that the classical literature does not. However, this flexibility comes at a price. It is not always clear whether such empirically founded comparative statics are primarily a reflection of the data on which the exercise is based, or if, like a 'theorem', they are a direct consequence of model assumptions. We proceed in the tradition of the Krugman model of economic geography and the Hecksher-Ohlin-Samuelson model of trade and examine a highly stylized geography. By doing this, we hope to develop insight into the way that the basic forces of urban economics – commuting costs, returns to scale and preference heterogeneity – interact to affect equilibrium, and hence to refine our understanding about which comparative statics are empirically ambiguous and which are not.

While a number of our findings are surprising, two are particularly noteworthy. The heterogenous preferences of the QSM framework imply that an average household must prefer central to peripheral work or residence. Such preferences have no analog in the classical literature and they encourage a concentration of work and residence in the central location even in the absence of any other agglomeration force. Second, although the set of interior equilibria is often unique, we find that corner equilibria exist whenever there are increasing returns to scale in production. Both interior and corner equilibria always occur as long as there are increasing returns. Our framework is similar enough to those typical of the QSM literature to suggest that these corner equilibria are a general feature of these models. Given that quantitative implementations of these models often depend on the uniqueness of equilibrium, we suspect this result may have important practical implications for the quantitative literature.

Determining the stability of equilibria is surprisingly subtle. The notion of stability most relevant to the quantitative literature, 'an iterative process will find the equilibrium', turns out not to be well defined. On the other hand, an analysis of a system of differential equations whose steady states coincide with the static equilibria of our model requires that we assume an ad hoc adjustment process and seems intractable. Given this, we rely on a more game theoretic notion of stability, in the spirit of trembling hand perfection. Using this definition, we show that: corner and 'near corner' equilibria are always unstable; for weak increasing returns to scale there is a unique, stable, interior equilibrium; and, in portions of the parameter space no stable equilibria exist.

Our benchmark model is based on a technology for which increasing returns to scale depends only on employment in a single location. We extend this model to a case where returns to scale can 'spillover' to nearby locations. Our results suggest that corner equilibrium do not persist with small productivity spillovers. They become 'near corner' equilibria that approach a corner as spillovers become weaker. More generally, the equilibria we find in our benchmark model appear to change continuously with the introduction of small spatial productivity spillovers.

Finally, for the cities of the Western world, we present a simple and brief 500 year history of the relationships between urban form and increasing returns to scale. We then compare the qualitative features of this history with the behavior of our model as increasing returns to scale varies over a range consistent with the approximately 13-fold increase income observed over this period. We find a qualitative relationship between the behavior we see in our model and history. Thus, our model appears to offer a simple theory that predicts the qualitative behavior of urban geography in the Western world from the pre-industrial period to the present.

## 1 Literature

We can usefully partition the literature into an older urban economics literature and a more recent literature on quantitative spatial models. Most papers in classical urban economics assume that households are homogenous or that there are a small number of types. Space is continuous and uniform, whether on a line or in a plane, and equilibrium cities are generally symmetric around a single exogenously selected point. The most influential model in this literature is the monocentric city model. This workhorse model rests on the assumption that the location of work is fixed exogenously at the center and households choose only their location of residence, although the model is otherwise quite general (Fujita, 1989).

The first general statement of our problem is due to Ogawa and Fujita (1980). This landmark paper considers a simple setting where firms choose only their location and households choose only their places of work and residence. They introduce the idea that firm productivity benefits from spillovers from every location, with distant spillovers less beneficial than those nearby. This assumption, now conventional, requires that the productivity of any given location responds to a distance weighted mean of employment at all locations.

This creates an agglomeration externality, while land scarcity acts as a dispersion force. As the benefits of spillovers increases relative to the cost of commuting, they observe first a uniform, then a duocentric, and finally a monocentric equilibrium. Lucas and Rossi-Hansberg (2002) revisits this problem and allow firms and households to substitute between labor and land and consumption and land. They establish general existence and uniqueness results, but otherwise rely on numerical methods and restrict attention to 'weak enough' increasing returns.

Recently, a second class of models has been brought to bear on problems of urban economics (Redding and Rossi-Hansberg, 2017). In the QSM literature, model cities consist of discrete sets of locations rather than continuous spaces, and they describe realistic rather than highly stylized geographies. More importantly, this literature considers heterogenous rather than homogenous agents. Where the older literature tends to focus on analytic solutions and qualitative results, the QSM literature focuses on the numerical evaluation of particular comparative statics in models that describe particular real world locations.

The QSM literature draws on a long history of scholarship on discrete choice models. Space is discrete and is described a matrix of pairwise commuting costs. These matrices are typically constructed to describe commuting costs between pairs of neighborhoods in the empirical application of interest. Households have heterogenous preferences over work-residence pairs and each household selects a unique pair. Locations are heterogenous in their amenities and productivity, and the possibility of endogenous agglomeration economies is sometimes considered.

In an important series of papers, Allen and Arkolakis (with coauthors) study the existence and uniqueness of equilibrium in spatial models similar to ours. Briefly, Allen and Arkolakis (2014) considers a model with a reduced form description of the land market. Land is entirely missing in Allen, Arkolakis and Takahashi (2020). Allen, Arkolakis and Li (2020) set the shares of residential and commercial land in each neighborhood exogenously. Allen, Arkolakis and Li (2015) is difficult to compare to our model because they "consider a general firm technology (e.g. it could be constant or decreasing returns to scale)" (p.4), which seems to rule out increasing returns to scale (see their equation (10)). While some of the Arkolakis and Allen results can probably be adapted to our setting, we do not investigate this possibility because our existence proof is straightforward and because our focus is on a characterization of multiple equilibrium, an issue that the Arkolakis and Allen theorems do not address.

## 2 A discrete city with heterogenous households

Our city consists of three locations $\mathcal{I} = \{-1,0,1\}$. This geography is the simplest in which to examine when activities concentrate in the center or disperse to the periphery. It is also qualitatively similar to the geography of the linear monocentric city. Each location is endowed with one unit of land. The city is populated by a continuum $[0,1]$ of households indexed by $\nu$ and by a competitive production sector whose size is endogenous. All households choose a residence $i \in \mathcal{I}$, a workplace $j \in \mathcal{I}$, and their consumption of housing and a tradable produced good.

In the QSM based literature, locations are typically endowed with both employment and residential 'amenities' that scale the payoffs from work and residence in each location. In the baseline model, we omit them for two reasons. First, it restricts our QSM based model to a stylized, featureless landscape of the sort considered in classical urban economics, and thereby facilitates the comparison of the two literatures. Second, it allows us to concentrate on implications of the fundamental economic forces of the model, commuting costs, returns to scale, and preference dispersion.

Each household $\nu \in [0,1]$ has a type $\mathbf{z}(\nu) \equiv (z_{ij}(\nu)) \in \mathbb{R}_+^{3\times 3}$, a vector of non-negative real numbers, one for each possible workplace-residence pair $ij$. The mapping $\mathbf{z}(\nu) : [0,1] \to \mathbb{R}_+^{3\times 3}$ is such that the distribution of types is the product measure of 9 identical Fréchet distributions:

$$F(\mathbf{z}) \equiv \exp\left(-\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{I}} z_{ij}^{-\varepsilon}\right). \tag{1}$$

Households have heterogenous preferences over workplace-residence pairs, and household types parameterize preferences. Thus, types describe preferences while $\varepsilon \in (0,\infty)$ describes the heterogeneity of preferences. An increase in $\varepsilon$ reduces preference heterogeneity and conversely.

Households commute between workplace and residence. Commuting from $i$ to $j$ involves an iceberg cost $\tau_{ij} \geq 1$. This cost is the same for all households and

$\tau_{ij} = 1$ if and only if $i = j$. Commuting costs affect household utility directly.[1]

A household that lives at $i$ and works at $j$ has an indirect utility

$$V_{ij}(\nu) = z_{ij}(\nu)\frac{W_j}{\tau_{ij}R_i^{\beta}}, \tag{2}$$

where $W_j$ is the wage paid at location $j$ and $R_i$ the land rent at $i$.

Let

$$S_{ij} = \left\{\mathbf{z} \in \mathbb{R}_+^{3\times 3}\,;\, V_{ij}(\mathbf{z}) = \max_{r,s\in\mathcal{I}} V_{rs}(\mathbf{z})\right\}$$

be the set of types $\mathbf{z}$ such that $ij$ is (weakly) preferred to all other location pairs $rs$. Then, using (1) and (2), the share $s_{ij}$ of households who choose the location pair $ij$ equals

$$s_{ij} = \mu\left(\mathbf{z}^{-1}(S_{ij})\right) = \frac{\left[W_j/(\tau_{ij}R_i^{\beta})\right]^{\varepsilon}}{\sum_{r\in\mathcal{I}}\sum_{s\in\mathcal{I}}\left[W_s/(\tau_{rs}R_r^{\beta})\right]^{\varepsilon}}, \tag{3}$$

where the last equality stems from the Fréchet distribution assumption and $\mu$ is the Lebesgue measure over [0,1].[2] Our model is static and so all choices occur simultaneously.

Because the $V_{ij}(\nu)$ are Fréchet distributed, the average utility $\overline{V}$ across all households equals:

$$\overline{V} \equiv \int_0^1 \max_{i,j\in\mathcal{I}} V_{ij}(\mathbf{z}(\nu))\mathrm{d}\nu = \mathit{\Gamma}\left(\frac{\varepsilon-1}{\varepsilon}\right)\left\{\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{I}}\left[W_j/\left(\tau_{ij}R_i^{\beta}\right)\right]^{\varepsilon}\right\}^{1/\varepsilon},$$

where $\mathit{\Gamma}(\cdot)$ is the gamma function. Households that share the same type choose the same location pair $ij$ and reach the same equilibrium utility level, while households who make the same choice may have different types and do not have the same equilibrium utility level. Likewise, households that choose different location pairs do not generally enjoy the same equilibrium utility level. In general, equilibrium utility varies with type.

We can rewrite (3) as,

$$s_{ij} = \kappa R_i^{-\beta\varepsilon}W_j^{\varepsilon}\tau_{ij}^{-\varepsilon}, \tag{4}$$

[1]These preferences are widely used in quantitative models, e.g., Ahlfeldt *et al.* (2015), Monte *et al.* (2018) and Heblich *et al.* (2020) to mention a few.

[2]Note that there might be types which belong to more than one $S_{ij}$ as they are indifferent between their multiple favorite choices. However, the set of such types always has a zero measure, so that they do not affect the type distribution (1).

for

$$\kappa \equiv \left[\Gamma\left(\frac{\varepsilon-1}{\varepsilon}\right)\right]^{\varepsilon}\overline{V}^{-\varepsilon}. \tag{5}$$

Let $M_i$ and $L_i$ be the mass of residents and households at $i$. Labor and land market clearing requires

$$\sum_{i\in\mathcal{I}} L_i = \sum_{j\in\mathcal{I}} M_j = 1, \tag{6}$$

where the residential population at $i$ is

$$M_i \equiv \sum_{j\in\mathcal{I}} s_{ij} = \kappa R_i^{-\beta\varepsilon} \sum_{j\in\mathcal{I}} W_j^{\varepsilon}\tau_{ij}^{-\varepsilon}, \tag{7}$$

and the labor force at $j$ is

$$L_j \equiv \sum_{i\in\mathcal{I}} s_{ij} = \kappa W_j^{\varepsilon} \sum_{i\in\mathcal{I}} R_i^{-\beta\varepsilon}\tau_{ij}^{-\varepsilon}. \tag{8}$$

Because each location $i$ is endowed with one unit of land, land market clearing requires

$$H_i + N_i \leq 1, \tag{9}$$

where $H_i$ is the amount of residential land and $N_i$ is the amount of commercial land at location $i$. Condition (9) implies that land rent is zero when the whole amount of land at $i$ is not used.[3]

Applying Roy's identity to (2), we have

$$H_i \equiv \sum_{j\in\mathcal{I}} s_{ij}H_{ij} = \sum_{j\in\mathcal{I}} s_{ij}\frac{\beta W_j}{R_i}. \tag{10}$$

Substituting (4) in (10) gives

$$R_i = \left(\frac{\beta\kappa}{H_i}\sum_{j\in\mathcal{I}} W_j^{1+\varepsilon}\tau_{ij}^{-\varepsilon}\right)^{\frac{1}{1+\beta\varepsilon}}. \tag{11}$$

This is the bid rent that households are willing to pay to reside at $i$, given $(W_1,...,W_I,H_i)$. Land rent accrues to absentee landlords who play no further role in the model.

Assume that the numéraire is produced under perfect competition and the production function at location $j$ is

$$Y_j = A_j L_j^{\alpha} N_j^{1-\alpha},$$

[3]To allow for corner equilibria, this condition should be written as $(H_i + N_i - 1)R_i = 0$.

where $A_j$ is location-specific TFP and $0 < \alpha < 1$. We assume that $A_j$ depends only on the level of employment at $j$,

$$A_j = L_j^{\gamma}, \tag{12}$$

for $\gamma \geq 0$, which is the same for all $j$. To keep our model as simple as possible, our description of returns to scale restricts attention to local scale economies when $\gamma > 0$. In Section 7, we consider the possibility of spatial externalities like those in Ogawa and Fujita (1980) and Lucas and Rossi-Hansberg (2002).

If location $j$ hosts a positive share of the production sector, the first-order conditions for the production sector yields the equilibrium wage and land rent as functions of land-labor ratio:

$$W_j = \alpha A_j \left(\frac{N_j}{L_j}\right)^{1-\alpha}, \tag{13}$$

$$R_j = (1-\alpha) A_j \left(\frac{L_j}{N_j}\right)^{\alpha}. \tag{14}$$

Dividing (13) by (14) and simplifying,

$$\frac{W_j}{R_j} = \frac{\alpha}{1-\alpha}\frac{N_j}{L_j}. \tag{15}$$

When $\gamma = 0$, note that

$$R_j = (1-\alpha)\left(\frac{N_j}{L_j}\right)^{-\alpha} = (1-\alpha)\left(\frac{W_j}{\alpha}\right)^{-\alpha/(1-\alpha)}. \tag{16}$$

This expression implies that $R_j$ and $W_j$ move in opposite directions under constant returns: a higher land rent at $j$ is equivalent to a lower wage at this location. Moreover, the land rent at $j$ is positive and finite if and only if $W_j$ is positive and finite.

Our city is described by $I$-vectors of real numbers. To describe them, define a *spatial pattern* to be an element of $\mathbb{R}^3_+$, a vector enumerating a non-negative real number for each location in $\mathcal{I}$. The spatial patterns that describe our model city are: the *residential pattern* $\mathbf{M}$; the *employment pattern* $\mathbf{L}$; the *housing pattern* $\mathbf{H}$; the *commercial pattern* $\mathbf{N}$; the *wage pattern* $\mathbf{W}$; and the *land rent pattern* $\mathbf{R}$. A spatial pattern is *interior* if none of its elements is zero. Otherwise, it is a *corner* pattern.

We can now define an equilibrium for our discrete city with heterogenous agents.

**Definition 1** *A spatial equilibrium is a vector* $\{\mathbf{M}^*,\mathbf{L}^*,\mathbf{H}^*,\mathbf{N}^*,\mathbf{W}^*,\mathbf{R}^*\}$ *such that: (i) all households make utility-maximizing choices of workplace, residence, land, and consumption to satisfy (3); (ii) the first-order conditions (13) and (14) are satisfied at every location where the tradable good is produced; (iii) the land market clearing condition (9) holds at each location; and (iv) the population constraint (6) holds.*

We say that an equilibrium is *interior* when all component patterns of $\{\mathbf{M}^*,\mathbf{L}^*,\mathbf{H}^*,\mathbf{N}^*,\mathbf{W}^*,\mathbf{R}^*\}$ are positive. Otherwise, it is a corner equilibrium. For tractability, and to ease comparison with the urban economics literature, we focus on *symmetric* spatial patterns $x = (x_{-1}, x_0, x_1)$ where $x_1 = x_{-1}$, and usually write such patterns as $(x_1, x_0, x_1)$ rather than $(x_{-1}, x_0, x_1)$. We say that a symmetric spatial pattern is *bell-shaped*, *flat*, or *U-shaped* as $x_0$ is greater than, equal to, or less than $x_1$. Say that a symmetric spatial pattern $(x_1, x_0, x_1)$ is *more centralized* than a symmetric spatial pattern $(y_1, y_0, y_1)$ if and only if $\frac{x_0}{x_1} > \frac{y_0}{y_1}$.

The restriction to three locations and symmetric patterns, allows us to focus our attention on the symmetric, three element versions of the spatial patterns for residence, employment, housing, industry, wages and land rents: $(M_1, M_0, M_1)$, $(L_1, L_0, L_1)$, $(H_1, H_0, H_1)$, $(N_1, N_0, N_1)$, $(W_1, W_0, W_1)$, and $(R_1, R_0, R_1)$.

Our analysis below is organized around studying the centrality of these six patterns. To facilitate this, define the centrality ratios,

$$m \equiv \frac{M_0}{M_1}; \quad \ell \equiv \frac{L_0}{L_1}; \quad h \equiv \frac{H_0}{H_1}; \quad n \equiv \frac{N_0}{N_1}; \quad w \equiv \frac{W_0}{W_1}; \quad r \equiv \frac{R_0}{R_1}.$$

Under symmetry, given an aggregate constraint, e.g. $2x_1 + x_0 = 1$, a spatial pattern is uniquely determined by one more piece of information, such as the ratio $x_0/x_1$.

For this linear city, the iceberg commuting cost matrix is

$$\begin{pmatrix} \tau_{-1,-1} & \tau_{-1,0} & \tau_{-1,1} \\ \tau_{0,-1} & \tau_{0,0} & \tau_{0,1} \\ \tau_{1,-1} & \tau_{1,0} & \tau_{1,1} \end{pmatrix} = \begin{pmatrix} 1 & \tau & \tau^2 \\ \tau & 1 & \tau \\ \tau^2 & \tau & 1 \end{pmatrix}, \tag{17}$$

where $\tau > 1$. To ease notation, we define

$$\begin{aligned} a &\equiv \frac{\alpha\beta}{1-\alpha} > 0, \\ b &\equiv \frac{(1-\alpha)(1+\varepsilon)}{\alpha\beta\varepsilon} = \frac{1+\varepsilon}{a\varepsilon}, \\ \phi &\equiv \tau^{-\varepsilon}. \end{aligned} \tag{18}$$

We will sometimes refer to $\phi \in (0{,}1)$ as the *spatial discount factor*. By inspecting (18), $\phi$ decreases with the level of commuting costs ($\tau \uparrow$) and increases with the heterogeneity of the population ($\varepsilon \downarrow$). Hence, $\phi$ may be high because either commuting costs are low, or the population is very heterogeneous, or both.

Finally, let

$$\omega \equiv w^{\varepsilon} \quad \text{and} \quad \rho \equiv r^{-\beta\varepsilon}. \tag{19}$$

Thus, $\omega > 1$ means that jobs at the central area pay a higher wage than those at the peripheries, while $\rho > 1$ means that land at the central area is cheaper than that at the peripheries.

Applying the symmetry assumption to the wage and land rent patterns **W** and **R** and using (4) and (17), we get the following equilibrium conditions:

$$\begin{aligned} s_{0,0} &= \kappa R_0^{-\beta\varepsilon} W_0^{\varepsilon}, \\ s_{0,1} = s_{0,-1} &= \kappa R_0^{-\beta\varepsilon} W_1^{\varepsilon} \tau^{-\varepsilon}, \\ s_{1,1} = s_{-1,-1} &= \kappa R_1^{-\beta\varepsilon} W_1^{\varepsilon}, \\ s_{1,-1} = s_{-1,1} &= \kappa R_1^{-\beta\varepsilon} W_1^{\varepsilon} \tau^{-2\varepsilon}, \\ s_{1,0} = s_{-1,0} &= \kappa R_1^{-\beta\varepsilon} W_0^{\varepsilon} \tau^{-\varepsilon}. \end{aligned} \tag{20}$$

We can now discuss the economic forces at work in this model. To illustrate ideas, consider the case when wages and land rents are the same in all locations and let $V = W/R^{\beta}$. In this case, using (2), a household's discrete choice problem is

$$\max_{ij} \left\{ \begin{array}{ccc} z_{-1,-1}V, & \frac{z_{-1,0}}{\tau}V, & \frac{z_{-1,1}}{\tau^2}V \\ \frac{z_{0,-1}}{\tau}V, & z_{0,0}V, & \frac{z_{0,1}}{\tau}V \\ \frac{z_{1,-1}}{\tau^2}V, & \frac{z_{1,0}}{\tau}V, & z_{1,1}V \end{array} \right\}.$$

This is the standard way of stating a discrete choice problem, except that we arrange the nine choices in a matrix so that the row choice corresponds to a choice of residence and choice of column to a choice workplace.

Because the distribution of idiosyncratic tastes is identical for all nine location pairs, the average payoff for a household choosing a central residence is

$$E\left(\max\left\{\frac{z_{0,-1}}{\tau}V, z_{0,0}V, \frac{z_{0,1}}{\tau}V\right\}\right) = \Gamma\left(\frac{\varepsilon-1}{\varepsilon}\right)\left(1+\frac{2}{\tau^{\varepsilon}}\right)^{1/\varepsilon} V. \tag{21}$$

Similarly, the average payoff for a household choosing one of the peripheral locations as a residence is

$$E\left(\max\left\{z_{-1,-1}V, \frac{z_{-1,0}}{\tau}V, \frac{z_{-1,1}}{\tau^2}V\right\}\right) = \Gamma\left(\frac{\varepsilon-1}{\varepsilon}\right)\left(1+\frac{1}{\tau^{\varepsilon}}+\frac{1}{\tau^{2\varepsilon}}\right)^{1/\varepsilon} V. \tag{22}$$

Because $\tau > 1$, it follows that the average payoff for a household choosing the peripheral location is lower than that of an average household choosing the central location. By symmetry, the corresponding statement is also true for the choice of work location. As a result, the average payoff for a household choosing a peripheral work location is lower than that of an average household choosing the central location for work. This discrete choice problem creates what we will call an *average preference* for residence in the central location, and a similar average preference for work in the center.

Because the average preference for central work and residence depends on both commuting costs, $\tau$, and population preference dispersion $\varepsilon$, it is not simply the translation into a discrete model of the commuting costs of classical urban economics. When $\varepsilon \to \infty$, (21) and (22) are identical, so that commuting costs alone are not sufficient to create an average preference for central work or residence. Nor are such average preferences an artifact of our stylized geography. If we exclude empirically uninteresting geographies like circles, most remaining location sets will have a center in the sense of this example, and corresponding average preferences for this location.

Average preferences for central work and location are noteworthy for the following reasons. First, even in the absence of more familiar agglomeration effects operating through production, this model has two 'agglomeration' forces, the average preference for central work and the average preference for central residence. Second, these average preferences are not agglomeration forces in the conventional sense. They do not incentivize geographic concentration, rather they incentivize concentration in the central location. Third, the urban economics literature based on homogenous agents takes seriously the possibility of differential labor productivity across locations. However, the possibility of preferences over work locations is never considered. This is a feature of quantitative spatial models without analog in the urban economics literature. Fourth, on the basis of the existing literature, we expect that the preference for central residence will be capitalized into higher central land rents. We will also see that the preference for central work location reduces central wages.

Against the two centralizing forces of average preferences are set two centrifugal forces. There is twice as much land in the periphery as the center. Because land contributes to utility and productivity, the scarcity of central land incentivizes the movement of employment and residence to the periphery.

We can now guess at the form of a symmetric equilibrium under constant returns to scale. In equilibrium, the centrifugal force of land scarcity and the centralizing force of average preferences balance. Access to the center will be scarce and so land rent will be higher and wages lower in the center. Whether the center ends up relatively specialized in residence or work will depend on which of the two activities has the highest demand for land, and this activity will locate disproportionately in the land abundant periphery.

We postpone a discussion of equilibrium under increasing returns to scale until after we provide a more formal characterization of such equilibria.

## 3 Existence and multiplicity of equilibria

We now turn to the characterization of symmetric equilibria in our three location city. We proceed in three main steps. In the first, we derive the demand for housing, the demand for commercial space, the supply of workers, and the supply of residents as functions of $\omega$ and $\rho$. This done, we derive a system of two equations in wages and land rent that characterize equilibrium. In the third step, we solve this system of equations.

The next proposition shows that in equilibrium all variables can be expressed in terms of just $\rho$ and $\omega$.

**Proposition 1** *The equilibrium demand for housing and commercial space and the equilibrium supply of workers and residents are:*

$$M_0 = \frac{\rho(\omega+2\phi)}{\rho(\omega+2\phi)+2(\phi\omega+1+\phi^2)}, \qquad M_1 = \frac{1-M_0}{2}, \tag{23}$$

$$L_0 = \frac{\omega(\rho+2\phi)}{\omega(\rho+2\phi)+2(\phi\rho+1+\phi^2)}, \qquad L_1 = \frac{1-L_0}{2}, \tag{24}$$

$$H_0 = \frac{a\rho(1+2\phi\omega^{-\frac{1+\varepsilon}{\varepsilon}})}{a\rho(1+2\phi\omega^{-\frac{1+\varepsilon}{\varepsilon}})+\rho+2\phi}, \qquad N_0 = 1-H_0, \tag{25}$$

$$H_1 = \frac{a(\phi\omega^{\frac{1+\varepsilon}{\varepsilon}}+1+\phi^2)}{a(\phi\omega^{\frac{1+\varepsilon}{\varepsilon}}+1+\phi^2)+\phi\rho+1+\phi^2}, \qquad N_1 = 1-H_1. \tag{26}$$

**Proof:** See Appendix A.

The derivation of these functions involves algebraic manipulation of the equilibrium conditions. For example, each of $M_i$ and $L_j$ is derived from the expressions for the share of households choosing each workplace-residence pair $ij$. In one case, we sum over workplace-residence pairs with a common residence, and in the other over pairs with a common workplace. The identical denominator for each of the four functions described by (23) and (24) is simply the three-location version of the denominator on the right-hand side of (3). The expressions for $H_i$ and $N_j$ are more complicated because they must also satisfy the land market clearing condition (9).

We can refine our understanding of how the average preference for central work and residence affect equilibrium by examining the expressions for residential population and employment (23)-(24). Assume that wages and land rents are equal across locations so that $\omega = \rho = 1$. Substituting in the expressions for $M_i$ and $L_i$ we find that $M_0(1) = L_0(1) > M_1(1) = L_1(1)$. Therefore, central employment and residence is greater than in the periphery even though the different locations have the same relative pecuniary appeal. This reflects the average preference for central residence and workplace.

To find the equilibrium values of $\omega$ and $\rho$ now derive a system of equations involving only $\omega$ and $\rho$ that incorporates all of the equilibrium conditions given in Definition 1.

**Proposition 2** *Assume $\gamma \neq \alpha/\varepsilon$. Then, a pair $(\rho^*,\omega^*)$ is an interior equilibrium if and only if it solves the following two equations:*

$$\omega^{\frac{1+\varepsilon}{\varepsilon}} = f(\rho) \equiv \frac{\phi\rho - 2a\phi\rho^{1+\frac{1}{\beta\varepsilon}} + (1+\phi^2)(1+a)}{(1+a)\rho^{1+\frac{1}{\beta\varepsilon}} + 2\phi\rho^{\frac{1}{\beta\varepsilon}} - a\phi}, \tag{27}$$

$$\omega^{\frac{1+\varepsilon}{\varepsilon}} = g(\rho;\gamma) \equiv \rho^{\frac{b}{1-\gamma\varepsilon/\alpha}} \left(\frac{\rho+2\phi}{\phi\rho+1+\phi^2}\right)^{\frac{\gamma\varepsilon/\alpha}{1-\gamma\varepsilon/\alpha}\frac{1+\varepsilon}{\varepsilon}}. \tag{28}$$

**Proof:** See Appendix B.

Combining (27) and (28), we arrive at a *single* equation in $\rho$ that determines the interior equilibria. Thus, studying the equilibrium behavior of our discrete linear city reduces to studying the solution(s) of one equation in one variable:

$$f(\rho) = g(\rho;\gamma). \tag{29}$$

We show the existence of an interior equilibrium by showing that (29) has a solution. We determine the number of possible interior equilibria by determining the number of such solutions of (29).

This argument requires two comments. First, (29) is not defined when $\gamma = \alpha/\varepsilon$. We will need a specific argument for this case. Second, we will see that $\gamma_m \equiv \alpha/\varepsilon$ is a threshold value of $\gamma$, below which there is a unique interior equilibrium, and above which multiple interior equilibria may occur.

We now turn to a characterization of equilibrium. To begin, we establish the properties of the functions $f$ and $g$. The following lemma states the main properties of function $f$ that are important for the characterization of spatial equilibria.

**Lemma 1** *There exist $\rho_0 \in (0,1)$ and $\rho_1 > 1$ such that $f(\rho)$ has a vertical asymptote at $\rho_0$ and is equal to 0 at $\rho_1$. Furthermore, $f(\rho) > 0$ if and only if $\rho_0 < \rho < \rho_1$ and decreases over $(\rho_0,\rho_1)$.*

**Proof:** See Appendix C.

Unlike $f$, the function $g$ varies with $\gamma$. We show in Lemma 2 that when $\gamma < \gamma_m$, $g$ is an increasing function that converges to an increasing step function as $\gamma \nearrow \gamma_m$. When $\gamma > \gamma_m$, $g$ is a decreasing function that converges to a decreasing step function as $\gamma \searrow \gamma_m$. In both limiting cases, the value $\rho_L$ at which the step occurs solves the equation

$$\rho^{\frac{1}{\alpha}} \frac{\rho + 2\phi}{1 + \phi\rho + \phi^2} = 1. \tag{30}$$

Because the left-hand side of (30) increases with $\rho$ and is equal to 0 (resp., $\infty$) when $\rho = 0$ (resp., $\rho \to \infty$), $\rho_L$ is unique and $\rho_L < 1$.

The following lemma provides a more formal statement of the relevant properties of $g$.

**Lemma 2** *(i) If $\gamma \neq \gamma_m$, then $g(\rho;\gamma)$ is strictly positive and finite over $[\rho_0,\rho_1]$. (ii) If $\gamma < \gamma_m$, then $g$ is increasing over $[\rho_0,\rho_1]$. (iii) If $\gamma > \gamma_m$, then $g$ is decreasing over $[\rho_0,\rho_1]$. (iv) As $\gamma$ converges to $\gamma_m$, we have:*

$$\lim_{\gamma \nearrow \gamma_m} g(\rho;\gamma) = \begin{cases} 0, & \rho < \rho_L; \\ \left(\frac{\rho_L + 2\phi}{1+\phi\rho_L+\phi^2}\right)^{-\frac{1+\varepsilon}{\varepsilon}} & \rho = \rho_L; \\ \infty, & \rho > \rho_L; \end{cases} \qquad \lim_{\gamma \searrow \gamma_m} g(\rho;\gamma) = \begin{cases} \infty, & \rho < \rho_L; \\ \left(\frac{\rho_L + 2\phi}{1+\phi\rho_L+\phi^2}\right)^{-\frac{1+\varepsilon}{\varepsilon}} & \rho = \rho_L; \\ 0, & \rho > \rho_L. \end{cases}.$$

**Proof:** See Appendix D.

Figure 1 illustrates $f$ and $g$. In each panel, the horizontal axis is $\rho$ and the vertical axis is an increasing transformation of $\omega$. The behavior of $f$, the red line in each panel, is relatively simple. It is a decreasing, continuous function that has a positive asymptote at $\rho_0 < 1$ and declines monotonically to 0 at $\rho_1 > 1$. The behavior of $g$ is more complicated. The panel on the left describes $g$ for three different values of $\gamma$, with dark blue the smallest, light blue the largest, medium blue in between, and all three less than $\gamma_m$. In every case, $g$ is a continuous, increasing function. The right panel is the same as the left, but considers larger values of $\gamma$. Here, the light blue line traces $g$ for the smallest value of $\gamma$, dark blue uses the largest value, medium blue is intermediate value, and all three are greater than $\gamma_m$.

This figure makes clear that, in general, $f$ and $g$ cross for a positive value $\rho$, and when $\gamma > \gamma_m$, they may cross more than once. Thus, an interior equilibrium exists in our model city throughout much or all of the parameter space.

Lemma 1 shows that $f$ has an asymptote at $\rho_0 < 1$ and is zero at $\rho_1 > 1$. Lemma 2 shows that as $\gamma \to \gamma_m$, $g$ also approaches its singularity at $\rho_L < 1$. While Lemmas 1 and 2 guarantee that $\rho_1 > \rho_L$, they do not allow us to order $\rho_L$ and $\rho_0$. Unsurprisingly, the equilibrium configuration of our city depends sensitively on whether or not $\rho_L > \rho_0$. Lemma 3 provides necessary and sufficient conditions on commuting costs and the demands for commercial and residential land for $\rho_L > \rho_0$.

**Lemma 3** *There exists a function $\overline{\phi}(\beta\varepsilon) \in (0,1)$ and scalar $\overline{a} > 0$ such that if $\phi < \overline{\phi}$ or $a < \overline{a}$ then $\rho_0 < \rho_L$. Conversely, if $\phi > \overline{\phi}$ and $a > \overline{a}$, then $\rho_0 > \rho_L$.*

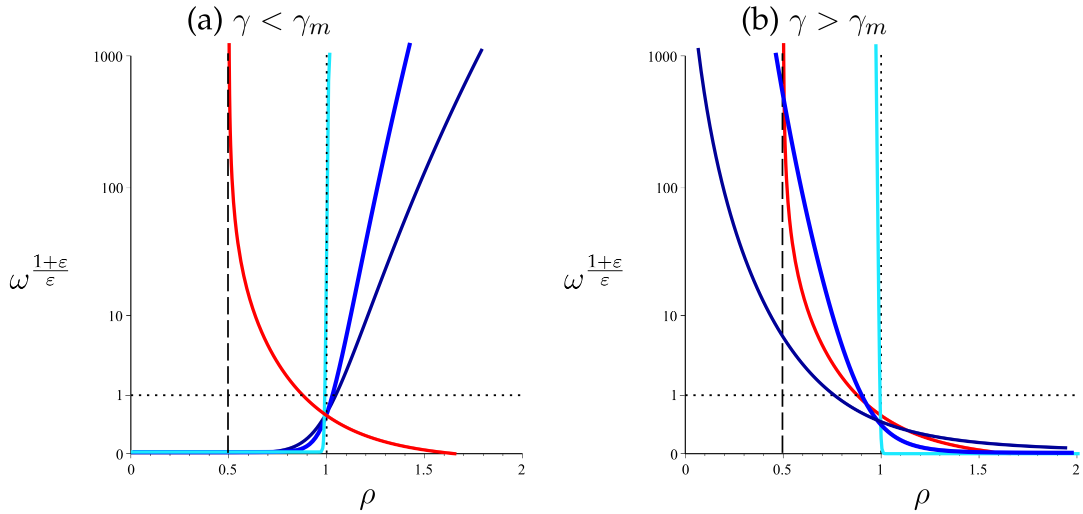
**Proof:** See Appendix E.

Restating this lemma informally, we have $\rho_0 < \rho_L$ if either the spatial discount factor is low *or* the demand for commercial land is sufficiently large relative to the demand for residential land. Conversely, if the spatial discount factor is high *and* the demand for commercial land is low, then $\rho_0 > \rho_L$.

The above equilibrium conditions imply the following relationship between land rent and employment, which will be useful in a number of our proofs. Using (24) leads to

$$\ell = \omega \frac{\rho + 2\phi}{\phi\rho + 1 + \phi^2}.$$

Figure 1: Graphical demonstration of equilibrium.




Notes: *In both panels, $f$ is the red line and $\rho_0 < \rho_L$. The blue lines describe $g$. In the left panel, darker colors of blue indicate smaller values of $\gamma$ and in the right panel darker colors of blue indicate larger values of $\gamma$.*

Combining this with (28), this equation becomes

$$\ell = \left( \rho^{\frac{1}{a}} \frac{\rho + 2\phi}{\phi\rho + 1 + \phi^2} \right)^{\frac{\gamma_m}{\gamma_m - \gamma}}. \tag{31}$$

If $\gamma < \gamma_m$, then $\ell$ increases over $(\rho_0,\rho_1)$ and $\ell(\rho) > 1$ if and only if $\rho > \rho_L$. On the other hand, when $\gamma > \gamma_m$, the opposite holds: $\ell$ decreases over $(\rho_0,\rho_1)$ and $\ell(\rho) > 1$ if and only if $\rho < \rho_L$.

So far, we have focused on interior equilibria. The following proposition formalizes this discussion and characterizes the corner equilibria.

**Proposition 3** *If $\gamma \geq 0$, an interior equilibrium always exists. Furthermore, there exist two corner equilibria if and only if $\gamma > 0$. These corner equilibria are such that the employment patterns are given, respectively, by $(0,1,0)$ and $(1/2,0,1/2)$. In both corner and interior equilibria, each location hosts a positive mass of residents.*

**Proof:** See Appendix F.

## 4 Comparative statics

We now turn our attention to an investigation of equilibrium changes as returns to scale, $\gamma$, commuting costs, $\tau$, and preference dispersion, $\varepsilon$ change.

## A *Returns to scale*

To ease exposition, we introduce terminology to describe the three important domains of returns to scale. These ranges will correspond to qualitatively different equilibrium behavior.

**Definition 2** *Increasing returns to scale (IRS) are: (i) weak if* $0 < \gamma < \gamma_m \equiv \alpha/\varepsilon$*; (ii) moderate if* $\gamma_m \leq \gamma \leq \gamma_s$*; or (iii) strong if* $\gamma > \gamma_s$*. Production is constant returns to scale when* $\gamma = 0$*.*

This definition requires three comments. First, population must be heterogeneous ($\varepsilon$ finite) for weak returns to scale to occur. Therefore, we cannot observe the weak increasing returns domain cities with the homogenous households that are standard in the urban economics literature. This establishes that qualitative features of the equilibrium depend on the degree of preference heterogeneity.

Second, we will see that the threshold level of returns to scale at which the possibility of discontinuous changes in equilibrium outcome most obviously arise is $\gamma_m = \alpha/\varepsilon$. In a modern economy, the labor share of production, $\alpha$, is about 0.6, while the range of commonly used estimates for $\varepsilon$ is about [5,7]. Taking the ratio of these values, we have $\gamma_m$ in [0.085,0.12]. Estimates of the wage elasticity of population for modern, developed country cities that control for sorting and first-nature productivity are typically around 0.05. However, the raw correlation between wages and density is larger, as are estimates for developing countries. This back of the envelope calculation, together with results presented below, will suggest that multiple interior equilibria occur in an empirically relevant part of the parameter space.

Third, $\rho_0$ must be smaller than $\rho_L$, i.e., commuting costs or the productivity of land sufficiently large, for moderate returns to scale to arise. To see this, solve the equilibrium condition $f(\rho) = g(\rho; \gamma)$ for $\gamma$ to get the function $\gamma(\rho)$.[4] Denoting by $\gamma_s$ the interior maximum of the function $\gamma(\rho)$ over $[\rho_0, \rho_1]$, it is readily verified that $\gamma_m = \gamma_s$ when $\rho_0 \geq \rho_L$, while $\gamma_m < \gamma_s$ when $\rho_0 < \rho_L$. In words, the region of moderate increasing returns to scale exists only when $\rho_0 < \rho_L$. As shown by Figure 2, $\gamma_s$ is the value where the equilibrium correspondence becomes singe valued.

[4]Specifically, we take the logarithm of $f(\rho) = g(\rho; \gamma)$ where $g(\rho; \gamma)$ is rewritten as (D.1). This yields a linear equation in $\gamma$, which has a unique non-negative solution $\gamma(\rho)$.

Figure 1 summarizes Lemmas 1-3 graphically when $\rho_0 < \rho_L$. Each panel evaluates $g$ for three different values of $\gamma$. In every panel, the light blue line describes $g$ for a value of $\gamma$ close to the weak/moderate threshold $\gamma_m$, and darker blue lines describe $g$ functions for values of $\gamma$ that are progressively further from this threshold. Taken together, these figures permit a fairly complete description of the interior equilibria of our linear city with three locations when $\rho_0 < \rho_L$.

The case of constant or weak increasing returns to scale is illustrated in panel (a) where high commuting costs encourage households to work where they live or land hungry production faces pressure to disperse to the periphery (or both). We expect an equilibrium in such an economy to exhibit low levels of commuting and dispersed production. In fact, regardless of $\gamma$, $f$ and $g$ cross slightly below $\rho = 1$ and at a moderate value of $\omega$. Since $\rho_0 < \rho_L$, as $\gamma \nearrow \gamma_m$ and $g$ approaches its asymptote at $\rho_L$, it does so when $f$ is well away from its asymptote at $\rho_0$. Hence, the equilibrium value of $\omega$ grows with $\gamma$ but remains bounded, meaning that the ratio of central to peripheral wages can never grow too large. Thus, panel (a) describes a city where neither employment nor residence are highly concentrated in either the center or periphery.

In panel (b), we first consider the case of moderate increasing returns described by the medium blue line. At this value of $\gamma$, $g$ crosses $f$ three times. At the first intersection point, we have $\rho_1^* < 1$ and $\omega_1^* > 1$; at the second one, we see that $\rho_2^*$ approaches $\rho_L$ as $\gamma$ decreases toward $\gamma_m$; at the third intersection point, we have $\rho_3^* > 1$ and $\omega_3^* < 1$. The value $\omega_1^*$ (resp., $\omega_3^*$) in turn requires that employment occurs primarily in the center (resp., splits between the two peripheral locations).

The light blue line in panel (b) describes $g$ when $\gamma$ is just above $\gamma_m$. As $\gamma$ approaches this threshold, for one of the two new equilibria $\omega$ grows without bound (and occurs outside the frame of the figure) while $\omega$ approaches zero in the other equilibrium. That is, just above the threshold, these two equilibria approach corner patterns where all employment is either central or peripheral. The dark blue line in panel (b) describes $g$ when returns to scale are strong. As $\gamma$ increases further, the remaining interior equilibrium involves a moderate value of $\omega$, that is, an equilibrium where employment is more or less evenly distributed across the three locations. Indeed, as households get wealthier, they consume more of the numéraire good and become more sensitive to the cost of commuting. As this occurs, they bid away differences in land rent and distribute themselves more uniformly across the three locations. Thus we arrive at the surprising conclusion that a low degree of increasing returns leads to agglomeration while a high degree fosters dispersion.

We now turn to a formal statement of our results. Propositions 4 to 7 confirms the intuition about equilibrium that we take from Figure 1 and extends it to the case where $\rho_0 > \rho_L$.

*a Constant returns to scale*

We start with the case where $\gamma = 0$. The following proposition characterizes the unique spatial equilibrium.

**Proposition 4** *Under constant returns to scale, there exists a unique equilibrium. This equilibrium is interior and such that* $0 < \rho^* < 1$ *and* $0 < \omega^* < 1$. *Furthermore, if*

$$\frac{\varepsilon}{1+\varepsilon} < \frac{1-\alpha}{\alpha\beta} \tag{32}$$

*holds, then the equilibrium employment pattern is bell-shaped.*

**Proof:** See Appendix G.

The inequalities $0 < \rho^* < 1$ and $0 < \omega^* < 1$ imply that all rents and wages are positive and finite. Furthermore, the equilibrium land rent is higher in the center while the equilibrium wage is lower, which reflects the average preference for the center. Hence, even in the absence of increasing returns, a (partial) agglomeration of production occurs at the center when (32) holds.

It is readily verified that, for $\gamma = 0$ the equilibrium condition (29) may be rewritten as (we use this condition twice)

$$\frac{\phi\rho - 2a\phi\rho^{1+\frac{1}{\beta\varepsilon}} + (1+\phi^2)(1+a)}{(1+a)\rho^{1+\frac{1}{\beta\varepsilon}} + 2\phi\rho^{\frac{1}{\beta\varepsilon}} - a\phi} = \rho^b. \tag{33}$$

When the population is homogeneous ($\varepsilon \to \infty$), the condition (33) has a unique solution $\rho^* = 1$, which in turn implies $\omega^* = 1$. Using (23) and (24) shows that $L_i^* = M_i^* = 1/3$ for $i = -1,0,1$, while it is easy to show that $s_{ij}^* = 0$ for $i \neq j$ and $s_{ii}^* = 1/3$. In other words, the spatial equilibrium is described by a flat world in which each location is an autarky. This should not come as a surprise as there is no agglomeration force.

*b Weak increasing returns*

The next two proposition describes the equilibrium when increasing returns to scale are weak.

**Proposition 5** *Assume that $0 < \gamma < \gamma_m$. Then, there is a unique interior equilibrium and $\rho^* < 1$. Furthermore, if (32) holds, then the equilibrium employment pattern is bell-shaped and such that*

$$\frac{d\ell^*}{d\gamma} > 0 \qquad \frac{d\rho^*}{d\gamma} < 0 < \frac{d\omega^*}{d\gamma}.$$

**Proof:** See Appendix H.

The substance of Proposition 5 seems intuitive. As scale economies increase, the central location attracts a larger share of workers, while both the relative land price and relative wage increase as the land rent and wage at the center capitalize the agglomeration force resulting from increasing scale economies.

The comparative statics in Proposition 5 hold whenever $0 < \gamma < \gamma_m$. More specifically, when $\gamma$ is slightly below $\gamma_m$, two cases may arise. (i) If $\rho_0 < \rho_L$, then setting $\gamma = \gamma_m$ in $[f(\rho)]^{1-\gamma\varepsilon/\alpha} = [g(\rho;\gamma)]^{1-\gamma\varepsilon/\alpha}$ yields an expression equivalent to (30) whose unique solution given by $\rho^* = \rho_L$. The equilibrium wage ratio remains bounded and the limiting employment pattern remains interior as $\gamma \nearrow \gamma_m$. (ii) If $\rho_0 > \rho_L$, then $\omega^* \to \infty$ and the equilibrium employment pattern converges to $(0,1,0)$ as $\gamma \nearrow \gamma_m$, hence $\rho^* = \rho_0$.

Less formally, there are two distinct equilibrium regimes. When commuting costs or land productivity is high, i.e. $\rho_0 < \rho_L$, the full concentration of production in the central location does not occur. The productivity advantage of the land abundant periphery is too great, or commuting is too costly to allow such central concentration of employment. On the other hand, when commuting costs are low and land is less productive, i.e. $\rho_0 > \rho_L$, then almost full concentration of employment at the CBD occurs. In this case, there is no interior equilibrium at $\gamma = \gamma_m$ because increasing returns are strong enough to complement the preference for central employment and generate almost full concentration of employment at the center.

*c Moderate increasing returns*

When $\gamma > \gamma_m$, the function $f$ remains unchanged, but the function $g$ changes from an increasing to a decreasing function. As both $f$ and $g$ are decreasing, we are no longer assured of the existence of a unique equilibrium.

The following proposition formalizes the intuition suggested by Figure 1 and extends it to the case where $\rho_0 > \rho_L$.

**Proposition 6** *Assume $\gamma$ is slightly above $\gamma_m$.*

*i. If $\rho_0 < \rho_L$, then there exist two interior equilibria, $(\rho_1^*,\omega_1^*)$ and $(\rho_3^*,\omega_3^*)$, such that, $\omega_1^* > 1 > \omega_3^*$ and $\rho_1^* < 1 < \rho_3^*$, as well as a third interior equilibrium. As $\gamma \searrow \gamma_m$, the first two equilibrium employment patterns converge to $(1/2,0,1/2)$ and $(0,1,0)$, while the third equilibrium is interior.*

*ii. If $\rho_0 > \rho_L$, then there exists a unique interior equilibrium $(\rho^*,\omega^*)$ such that, $\omega^* < 1$ and $\rho^* > 1$. Furthermore, as $\gamma \searrow \gamma_m$, the equilibrium employment pattern converges to $(1/2,0,1/2)$.*

**Proof:** See Appendix I.

Part (i) of Proposition 6 establishes the existence of multiple interior equilibria when either commuting costs are high or the productivity of land in production is relatively low.

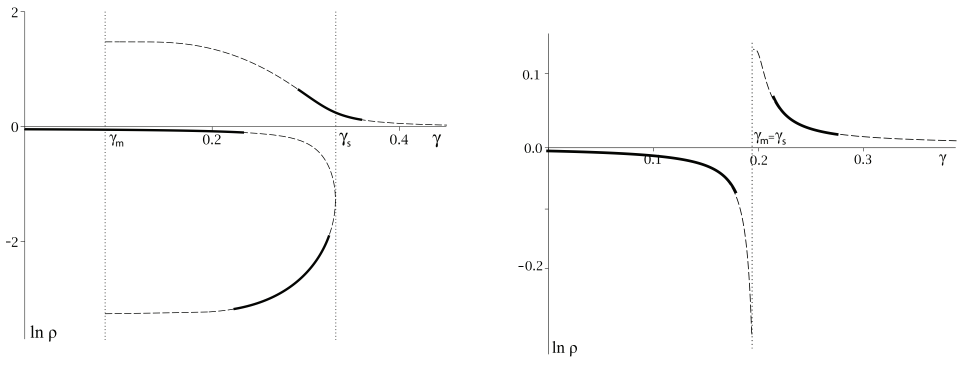
Proposition 6 characterizes equilibrium just above the threshold separating weak and moderate increasing returns, $\gamma_m$, while Proposition 7 characterizes equilibrium above $\gamma_s > \gamma_m$, the threshold separating moderate and strong increasing returns. It is natural to expect that the behavior we observe near $\gamma_m$ persists throughout the full range of moderate increasing returns, $(\gamma_m,\gamma_s)$. We demonstrate that this is sometimes the case in the examples illustrated in Figure 2. In fact, we cannot rule out the possibility of more complicated equilibrium behavior for values of $\gamma$ just below $\gamma_s$, although we cannot find a counter example to contradict the conjecture that the results of Proposition 6 hold throughout the range of moderate increasing returns.

*d Strong increasing returns*

Figure 1 shows that only a single interior equilibrium persists when $\gamma$ exceeds $\gamma_s$. More generally, we have:

Figure 2: Equilibrium correspondence between $\rho$ and $\gamma$.

$\rho_0 < \rho_L$ $\rho_L < \rho_0$




Notes *In both panels the $x$-axis is $\gamma$ and the $y$-axis is* $\ln \rho$*. The left panel illustrates all interior equilibria as $\gamma$ varies when $\rho_0 < \rho_L$. The right panel shows the case where $\rho_L < \rho_0$. Solid lines indicate stable equilibria and dashed lines indicate unstable equilibria, where stability is defined as in Section 5.*

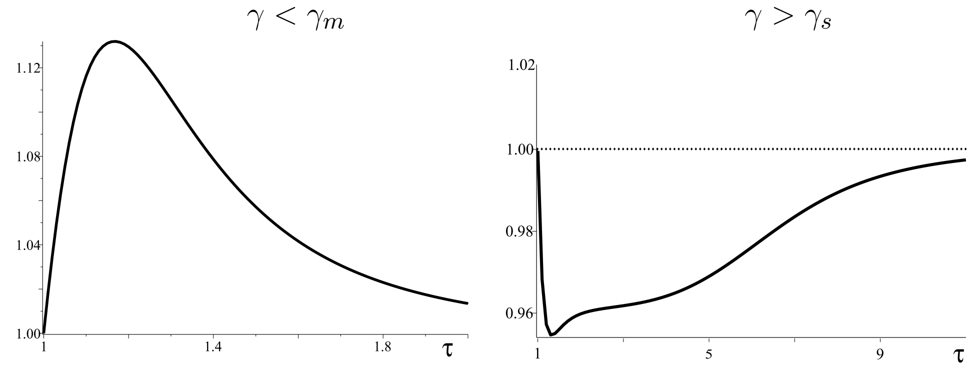
**Proposition 7** *If $\gamma > \frac{1+\varepsilon}{(1-\beta)\varepsilon} - \alpha \geq \gamma_s$, then there exists a unique interior equilibrium. Furthermore, the equilibrium employment pattern gets flatter as $\gamma$ rises and converges to the uniform pattern when $\gamma \to \infty$.*

**Proof:** See Appendix J.

To our knowledge, Proposition 7 is new to the literature. While it has long been understood that increasing returns to scale could lead to multiple equilibria, the idea that sufficiently high increasing returns leads, once again, to a unique equilibrium is novel. Even more surprising, this equilibrium involves more dispersion of production as the degree of increasing returns rises. The intuition behind this result seems straightforward. As $\gamma$ increases sufficiently, all else equal, wages must rise. This makes households better-off, which in turn increases the utility loss from commuting. As the marginal utility of income falls, households solve this trade-off by reducing the total amount of commuting, which leads to more dispersion.

Figure 2 describes the equilibrium correspondence between $\gamma$ and $\rho$ for numerical examples satisfying $\rho_0 < \rho_L$ in panel (a) and $\rho_L < \rho_0$ in panel (b). In both panels, the $x$-axis describes $\gamma$ and the $y$-axis describes $\ln \rho$. Both figures show all interior equilibria, but not the corner equilibria required by Proposition 3. Both figures anticipate our analysis of stability in Section 5 and indicate stable equilibria with a solid line and unstable equilibria with a dashed line.

Figure 3: Employment and residence ratios as commuting cost varies.




Notes: *In both panels, the $x$-axis is $\tau$ and the heavy black line shows the equilibrium value of $\ell = L_0/L_1$. In panel (a) $\gamma = 0.07$ and in panel (b) $\gamma = 0.7$. Other parameters used for these calculations are $\alpha = 0.9, \beta = 0.25, \varepsilon = 5$.*

### B *Commuting cost and preference dispersion*

We now consider how interior equilibria vary as commuting cost or preference dispersion change, with a special emphasis on the cases of a homogeneous or infinitely heterogeneous population when $\gamma > 0$. Both panels of Figure 3 illustrate the equilibrium employment ratio $\ell = L_0/L_1$ as $\tau$ varies. The left panel illustrates an example where returns to scale are weak, $\gamma < \gamma_m$, and the right panel an example where they are strong, $\gamma > \gamma_s$. We omit the intermediate case to avoid the complication of multiple equilibria.

In the case of weak increasing returns, the concentration of employment in the center is increasing in commute costs for low levels of $\tau$ and decreasing for high levels of $\tau$. Equilibrium employment is uniform as $\tau$ approaches either one or infinity. The peak of the employment ratio locus occurs around $\tau = 1.2$, where commuting results in a 20% utility penalty. For reference, Redding and Turner (2015) report an average round trip commute of about 50 minutes for an average American, this is about 12% of an eight hour work day. This calculation suggests that the complicated comparative statics illustrated in panel (a) could well be empirically relevant.

The case of strong returns to scale illustrated in panel (b) is about opposite. For $\tau$ to the left of the minimum of the employment ratio curve, the concentration of employment in the periphery increases as commute costs increase. To the right of the peak, this behavior reverses. Like the case of weak increasing returns illustrated in panel (a), as $\tau$ approaches one or infinity, the employment pattern becomes flat. Note that the scale of the $x$-axis in panel (b) is different from that in panel (a) and the peak of the employment ratio curve occurs around $\tau = 2$. The empirically relevant region of panel (b) likely lies to the left of the minimum of the employment ratio curve. In this region, increases in $\tau$ decentralize employment.

Recalling that $\phi = \tau^{-\varepsilon}$, it is easy to see that $\phi \to 1$ when $\tau \to 1$ or $\varepsilon \to 0$, while $\phi \to 0$ when $\tau \to \infty$ or $\varepsilon \to \infty$. Equations (21) and (22) show that the average payoff of the choice of central versus peripheral workplace converge toward each other when $\phi$ goes to either zero or one. This is consistent with the convergence to a flat equilibrium for extreme values of $\tau$ that we see in Figure 3 and suggests that such convergence may be general.

To show that this, taking the limit for $\phi \to 0$ and $\phi \to 1$ in (27) and (28), we obtain:

$$\begin{aligned}
f(\rho)|_{\phi=0} &= \rho^{-1-\frac{1}{\beta\varepsilon}}, \qquad f(\rho)|_{\phi=1} = \frac{\rho - 2a\rho^{1+\frac{1}{\beta\varepsilon}} + 2(1+a)}{(1+a)\rho^{1+\frac{1}{\beta\varepsilon}} + 2\rho^{\frac{1}{\beta\varepsilon}} - a}, \\
g(\rho;\gamma)|_{\phi=0} &= \rho^{\frac{1}{a}+\frac{\gamma}{\alpha\varepsilon}} \rho^{\frac{\gamma\varepsilon/\alpha}{1-\gamma\varepsilon/\alpha}\frac{1+\varepsilon}{\varepsilon}}, \qquad g(\rho;\gamma)|_{\phi=1} = \rho^{\frac{b}{1-\gamma\varepsilon/\alpha}}.
\end{aligned}$$

Evaluating each of these functions at $\rho = 1$ shows that $f(1)|_{\phi=0} = g(1;\gamma)|_{\phi=0} = 1$ and $f(1)|_{\phi=1} = g(1;\gamma)|_{\phi=1} = 1$. Hence, in both cases $\rho^* = 1$ is an interior equilibrium.

As $\varepsilon \to \infty$, preference dispersion disappears and households become homogenous. Because the domain of weak returns to scale vanishes when $\varepsilon \to \infty$, only the cases of moderate or strong increasing returns are relevant. The distribution of activity is also uniform when returns to scale are strong. Indeed, letting $\varepsilon \to \infty$ in $f(\rho)|_{\phi=0} = g(\rho;\gamma)|_{\phi=0}$, we have $\rho^{-1} = \rho^{\frac{1}{a}-1}$ whose unique interior solution is $\rho^* = 1$. It then follows from (27) that $\omega^* = 1$, which implies $\ell^*(1) = m^*(1) = h^*(1) = n^*(1) = 1$. Using (23) and (24), it is easy to show that $s^*_{ij} = 0$ for $i \neq j$ and $s^*_{ii} = 1/3$. Thus, when households no longer care about where they live and work, they focus on their own consumption only. In this case, high increasing returns make households rich enough that they prefer to work where they live at a lower wage in order to avoid larger utility losses from commuting. We may appeal to proposition 7 to say that the flat pattern is the only interior equilibrium when $\gamma > 1/(1-\beta) - \alpha$. In spite of strong increasing returns to scale, when the households are homogeneous the only interior equilibrium involves a uniform distribution of production activities across locations. *In sum, the equilibrium city consists of backyard capitalists who work where they live.*

However, it is possible that the limit does not describe what happens when agents are perfectly homogenous because the c.d.f. of the Fréchet distribution is a step function at $z = 0$. Assuming from the outset that households are homogeneous, that is, the indirect utility of a $ij$-household is given by $V_{ij} = W_j/(\tau_{ij} R_i^{\beta})$, we can show that the equilibrium involves three autarkic locations.[5] This result resembles Starrett's (1978) Spatial Impossibility Theorem. This theorem states that, regardless of the technology (decreasing, constant or increasing returns), in the absence of first nature any equilibrium (if it exists) in the Arrow-Debreu model of general equilibrium must be such that each location is an autarky. By contrast, if we allow for spatial spillovers across locations like in Osaka and Fujita (1980) and Lucas and Rossi-Hansberg (2002), agglomeration may arise.

As $\varepsilon \to 0$, we see that $\gamma_m \to \infty$, which means that returns to scale must be weak. Therefore, Proposition 5 applies and the interior equilibrium is unique. Interestingly, the flat pattern also emerges as the population becomes infinitely heterogeneous. Indeed, as $\varepsilon \to 0$, taste heterogeneity over pairwise choices becomes increasingly important relative to commuting costs and, in the limit, households ignore land price and wage differences, and the distribution of households across pairs is uniform. Likewise, as $\tau \to \infty$ the cost of commuting grows so high that it never makes sense to commute. Since the distribution of types is the same across locations, households must be uniformly distributed across locations. In this case, only autarky is possible and all households work where they live. In both cases, the distribution of residence and employment is uniform, but the city functions differently in the two cases. When $\varepsilon \to 0$, we have an extreme form of urban sprawl where many people commute while there is no city center. By contrast, when $\tau \to \infty$ we have a city of backyard capitalists. Formally, when $\varepsilon \to 0$, the payoffs for each of the nine location pairs become identical. As $\varepsilon \to \infty$, the payoff (22) attached to off-diagonal pairs goes to zero.

[5]The proof is available from the authors upon request.

In the former case, we have symmetric cross-commuting between any location pair, while we have no commuting at all in the latter case.

The following proposition summarizes this discussion.

**Proposition 8** *Interior equilibrium employment, residential, housing, and commercial patterns converge to a flat pattern when one of the following conditions holds: (i)* $\tau \to 1$ *or* $\tau \to \infty$ *and* $\gamma < \gamma_m$ *or* $\gamma > \frac{1+\varepsilon}{(1-\beta)\varepsilon} - \alpha$*; (ii)* $\varepsilon \to 0$*; or (iii)* $\varepsilon \to \infty$ *and* $\gamma > \frac{1}{1-\beta} - \alpha$*.*

We regard these comparative statics as surprising and important for two reasons. First, in the monocentric city model, decreases in commuting costs lead households to spread out. Our results contradict this intuition. In our heterogenous household model, comparative statics on commuting costs are not monotone. Second, as we describe in our review of the literature, discrete heterogenous agent models, similar to ours, are the basis for a rapidly growing quantitative literature. Often, such quantitative exercises evaluate the effects of counterfactual changes in commuting costs. To the extent that these counterfactual exercises are comparative statics of commuting costs, our results suggest the qualitative features of such counterfactual exercises may change sign in response to changes in incidental parameters.

# 5 Stability

It is common to appeal to stability as a selection device in the presence of multiple equilibria. This leads to the question of how to define stability. One candidate, particularly relevant for the quantitative literature, is to say that an equilibrium is stable if an iterative process will converge to it. Formally, if equilibria are defined by $f(\rho) = g(\rho)$ then equilibria are fixed points of $h(\rho) = \rho$, for $h(\rho) \equiv f^{-1}(g(\rho))$. It is well known that an iterative process will find a fixed point $\rho^*$ if and only if $|h'(\rho^*)| < 1$. Surprisingly, this notion of stability is not well defined. To understand the problem with this definition, observe that $h(\rho) = \theta\rho + (1-\theta)\rho$ also defines solutions of $f(\rho) = g(\rho)$, and therefore, that fixed points of $\widetilde{h}(\rho) = [(h(\rho) - (1-\theta)\rho)/\theta] = \rho$ are also solutions of $f(\rho) = g(\rho)$. However, the stability properties of this second equation may be different from the original. By choosing $\theta$ sufficiently small, we guarantee that $|\widetilde{h}'(\rho^*)| > 1$. Thus, an iterative notion of stability is not invariant to different, equivalent ways of formulating the equilibrium conditions.

Symmetry implies that we must determine the values of three variables to obtain the equilibrium outcome. For example, it is sufficient to know $L_0$, $M_0$, and $s_{00}$ to determine all the $s_{ij}$, and hence the vector $\{$**M**,**L**,**H**,**N**,**W**,**R**$\}$. Therefore, a second approach to stability requires that we specify an adjustment process describing how $L_0$, $M_0$ and $s_{00}$ respond to a perturbation. Stability is then well defined in the resulting dynamic system. This approach is subject to two problems. First, it is likely to be intractable. Second, it must rest on ad hoc descriptions of the adjustment process, and we suspect that the stability of any particular equilibrium is likely to be sensitive to these assumptions.

These difficulties lead us to a more game-theoretic notion of stability. In the spirit of trembling hand perfection, we say that an equilibrium is stable if households want to return to the equilibrium when an arbitrarily small measure of them are displaced. This definition of stability has three advantages. First, like our model, it is static and does not require an explicit description of time. Second, and unlike the other candidate definitions of stability, it has explicit behavioral foundations. Third, as we will see, it is tractable.

Let $ij$ and $kl$ be two arbitrary location pairs; $ij = kl$ (location pairs are equal) when $i = k$ and $j = l$ hold simultaneously and distinct otherwise. We say that an equilibrium is *unstable* if, for some $ij \neq jk$, for any arbitrarily small $\Delta > 0$, there is a subset of individuals of mass $\Delta$ who strictly prefer the location pair $kl$, which differs from their utility-maximizing pair $ij$, when a perturbation moves them all to $kl$. In other words, the subset of individuals who have been moved away from $ij$ do not want to move back. Otherwise, the equilibrium is *stable*.

The key issue is to determine the subset of individuals to use to check whether the equilibrium is unstable. In what follows, we assume that this subset is formed by individuals whose types are close to those of an individual indifferent between her equilibrium pair $ij$ and another location pair $kl$. Lemma 4 establishes that such an individual always exists.

Consider an equilibrium commuting pattern $\mathbf{s}^*\equiv(s^*_{ij})$, which could be interior or corner, and two location pairs, $ij$ and $kl$, such that $ij \neq kl$ and $s^*_{ij} > 0$. We say that an individual $\nu \in [0,1]$ is *indifferent between $ij$ and $kl$* if and only if

$$V^*_{ij}(\nu) = V^*_{kl}(\nu) \geq V^*_{od}(\nu), \tag{34}$$

for every location pair $od$ such that $od \neq ij$ and $od \neq kl$. Given this definition, we have:

**Lemma 4** *For any two distinct location pairs $ij$ and $kl$ such that $s_{ij}^* > 0$, there exists an individual $\nu \in [0,1]$ with $z_{ij}(\nu) \in S_{ij}$ and $z_{kl}(\nu) > 0$ who is indifferent between $ij$ and $kl$.*

**Proof:** See Appendix K.

**Definition 3** *Consider an arbitrarily small subset of individuals of measure $\Delta > 0$ who choose $ij$ and have types close to $\mathbf{z}(\nu) \in S_{ij}$ where $\nu$ is indifferent between $ij$ and $kl \neq ij$. If this individual is strictly better off when she and her neighboring individuals are relocated from $ij$ to $kl$, the spatial equilibrium is unstable. Otherwise, the spatial equilibrium is stable.*

The motivation for this definition is as follows. If the relocation of a small group of almost indifferent individuals from $ij$ to $kl$ makes the indifferent agent strictly better off, then, by continuity there is a non-negligible subset of individuals who strictly prefer $kl$ to $ij$. Hence, these individuals will never switch back to $ij$. On the contrary, if the indifferent individual never becomes strictly better off for any small subset, no other individual strictly prefers a different location pair. Hence, all the individuals will be willing to switch back to $ij$.

By relocating a small subset of individuals from $ij$ to $kl$, the commuting pattern $\mathbf{s}$ becomes different from the equilibrium pattern $\mathbf{s}^*$. Hence, for our definition of stability to make sense, we must be able to compare the equilibrium and off-equilibrium utility levels. For this to be possible, we must determine the conditional equilibrium vectors of wages and land rents $\overline{\mathbf{W}}(\mathbf{s})$ and $\overline{\mathbf{R}}(\mathbf{s})$. We show in Appendix L, that, for $\alpha > 1/2$, these vectors exist, are unique and continuous in $\mathbf{s}$.

Using the above definition of stability, we are now equipped to study the stability of the equilibria identified in Proposition 9. We start with corner equilibria.

**Proposition 9** *The corner equilibria are always unstable.*

**Proof:** See Appendix L.

This result is easy to understand. Consider the agglomerated corner equilibrium $\mathbf{L}^* = (0,1,0)$. No single individual wants to move to, say, location 1 because her marginal productivity would be zero. This is why $(0{,}1{,}0)$ is an equilibrium employment pattern. By contrast, when a small subset of workers happens to be at 1, (13) implies the income of those whose tastes are very close to those of the indifferent individual is extremely high. As a consequence, they do not want to move back to location 0.

If we rely on stability to select among multiple equilibria, it means that we can ignore the corner equilibria. This result has the potential to greatly simplify quantitative exercises based on this family of models.

By Proposition 3, interior equilibria always exist. The next proposition provides necessary and sufficient conditions for an interior equilibrium to be stable. Denote by $\rho_{CR}$ be the equilibrium that prevails under constant returns.

**Proposition 10** *There exists a function $\mathbb{F}(\rho)$ independent of $\gamma$ such that an interior equilibrium $\rho^*$ is stable if and only if $\mathbb{F}(\rho^*) > 1$. This function is continuous over $(\rho_0,\rho_{CR})$ and over $(\rho_{CR},\rho_1)$, satisfies $\mathbb{F}(\rho_0) = \mathbb{F}(\rho_1) = 0$, and has a vertical asymptote at $\rho = \rho_{CR}$.*

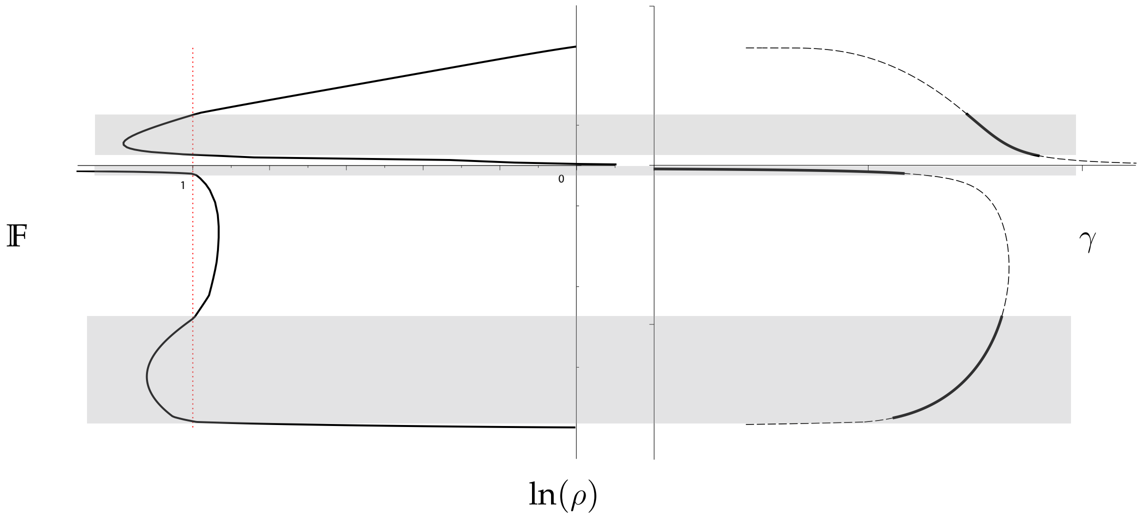
**Proof**: See Appendix M where we also give the explicit form of function $\mathbb{F}$.

Proposition 10 provides a simple test for checking the stability of any interior equilibria. In fact, it also allows us to draw general conclusions about stability. The left panel of Figure 4 illustrates the function $\mathbb{F}$. In this panel, the value of $\mathbb{F}$ is on the $x$-axis and $\ln \rho$ on the $y$-axis (strictly, this is a plot of the correspondence $\mathbb{F}^{-1}$). The critical value $\mathbb{F} = 1$ is indicated by the dotted vertical line. As in Figure 2, we transform the $y$-axis for legibility. From this figure, we can read off the intervals of $\rho$ where the equilibrium is stable by checking whether $\mathbb{F} > 1$. These regions are indicated by the horizontal shaded bands.

In the right panel of Figure 4 we repeat the left panel of Figure 2. This figure describes the equilibrium correspondence between $\gamma$ and $\ln \rho$. Since the $y$-axes on the two panels agree, we can put them next to each other and project the stable intervals onto the equilibrium correspondence between $\gamma$ and $\ln \rho$ given in the right panel. This done, we can also read off the intervals of $\gamma$ for which some or all of the possible interior equilibria are stable. As in Figure 2, the heavy lines indicate stable equilibria, and dashed lines indicate unstable equilibria.

The fact that $\mathbb{F}$ always has a vertical asymptote at $\gamma = 0$ guarantees that for $\gamma$ sufficiently close to zero, equilibrium will always be stable. Furthermore, from

Figure 4: Graphical representation of Proposition 10.




Notes: *The right half of this figure plots the equilibrium correspondence between* $\ln(\rho)$ *and* $\gamma$ *reproduced from figure 2. The* $x$*-axis in this half of the figure is* $\gamma$*. As in figure 2, heavy lines indicate stable equilibria and dashed lines indicate unstable equilibria. The left half of this figure uses the same parameter values to plot the function* $\mathbb{F}$ *(times minus one) from Proposition 10 on the* $x$*-axis and* $\ln(\rho)$ *on the* $y$*-axis. From Proposition 10,* $\rho$ *is a stable equilibrium whenever* $\mathbb{F} > 1$*. The* $\mathbb{F} = 1$ *threshold is illustrated by the vertical dashed line in the left panel. Intervals of* $\ln(\rho)$ *where* $\mathbb{F} > 1$ *and equilibrium is stable are indicated by the shaded horizontal bands.*

our discussion of equilibrium in the weak IRS case, we know that $\rho^*$ is close to $\rho_0$ or $\rho_1$ if and only if the equilibrium is close enough to a corner equilibrium. Therefore, because $\mathbb{F}(\rho_0) = \mathbb{F}(\rho_1) = 0$, it must be that these equilibria are unstable.

## 6 A very short introduction to European urban history

Explaining the evolution of cities and their role in the process of economic growth over the course of history has long been central to urban economics. In what follows, we examine how an equilibrium city changes as the degree of increasing returns, $\gamma$, increases and compare these changes to the history of European cities from the pre-industrial period to the present.

The historical record, surveyed below, indicates that returns to scale have increased over the past several centuries. The evolution of urban density during this time has been non-monotonic. Cities formed and became denser and employment more concentrated, before gradually becoming less dense with employment less concentrated. As Figure 2 demonstrates, this basic pattern of increasing, and then decreasing density and concentration of employment can be observed along every one of the possible equilibrium paths predicted by our model as returns to scale increase. Thus, our model rationalizes the joint evolution of returns to scale and these basic features of cities.

To focus on returns to scale, we ignore other important trends, e.g., decreases in transportation cost and in the land intensity of production. We have three reasons for this. First, up until now, the theoretical understanding of the implications of changes in returns to scale has been restricted by the intractability of this problem. This is where our results are the most novel. Second, as we will see, changes in returns to scale alone can explain qualitative features of the history of cities. Finally, as important as other changes have been, increases in productivity are surely more important. By focusing on changes returns to scale we keep attention on the central economic force behind the industrial revolution.

The existence of various types of agglomeration effects in modern cities is well established, even if debate continues about exactly how large they are. We do not have systematic evidence about time series change in the importance of agglomeration effects, but three well-established facts from economic history suggests that returns to scale are increasing. First, during the period around 1500, de Vries (1984) reports 154 European cities with population above 10,000, while Bairoch (1988) reports 89 cities of at least 20,000. At this time, Europe (without Russia) had only between 10 and 12 cities of more than 100,000 inhabitants. By 1800, the count of cities with a population of at least 10,000 and 20,000 increased to 364 and 194, respectively. Similarly, the share of the urban population was low and rose slowly from 10.7 in 1500 to 12.2 percent in 1750. The urbanization rate was still around 12 percent in 1800, but grew rapidly to 19 percent in 1850, 38 percent in 1900, 51 percent in 1950, and 75 percent in 2000 (Bairoch, 1988). Before the industrial revolution, Europe (without Russia) has no city of more than 2 million, but by 1910, four European cities cross this threshold (Berlin, London, Paris, and Vienna). Without a doubt, European cities have become progressively more attractive. While these increases are surely not entirely due to increases in returns to scale in production, it is equally sure that stronger returns to scale are partly responsible.

Second, economic historians have documented many small changes in pre-industrial Europe that contributed to the productivity of cities. For example, Cantoni and Yuchtman (2014) observe the spread of Universities in 14$^{\text{th}}$ century Germany and argue that, by spreading knowledge of Roman law, these Universities contributed to larger and more productive market cities. Dittmar (2011) documents the spread of the printing press and finds that it contributed causally to the sizes of cities where it was introduced. Finally, de la Croix *et al.* (2018) argue that apprenticeship may have played a similar role by facilitating the transmission of tacit knowledge. In all, the history of the pre-industrial period points to slow increases in the productivity of cities.

Finally, the nature of production has changed in a way that indicates an increase in returns to scale. For the period prior to the industrial revolution, de Vries (1984) reports on the pervasiveness of proto-industrialization in manufacturing. Under this system, rural households performed manufacturing at home, often of textiles, using materials provided by manufacturers. It is hard to imagine a system of industrial production that more strongly suggests constant or decreasing returns to scale in manufacturing.

Summing up, we can be confident that agglomeration and scale economies are positive in the modern economy and the available evidence strongly suggests that they were small and increasing in the pre-industrial period. Taken together these two observations indicate a trend upwards.

We now turn to a stylized description of the evolution of cities from the pre-industrial period to the present. de Vries (1984) describes cities in pre-industrial Europe as being organized as much for protection as production. Much of the population was employed in agriculture, either within city walls or without, manufacturing was at least as rural as urban, and the preponderance of urban residents were abjectly poor. While the absolute poverty of an average urban resident is clear, there is suggestive evidence from the US for the relative poverty of urban residents as well. Costa (1984) reports in a survey of US soldiers in the US civil war, early in the US industrial revolution, and finds that soldiers with urban backgrounds were shorter than those from rural backgrounds.

Clark (1951) is the first effort to provide systematic evidence about patterns of urban density. In this landmark study, Clark collects historical census data for twenty US and European cities. While the precise time period he considers varies from city to city, his data often begins early in the industrial revolution and continues until the early 20th century. His findings are unequivocal: center city density falls and density gradients flatten over time.

Turning to the late 20th and early 21st centuries, Baum-Snow (2007) documents the decline in US central city population from 1950 to 1990. Glaeser and Kahn (2004) find that in late 20th century US, employment follows decentralizing population to the suburbs and that peripheral urban residents of US cities tended to have much longer commutes than did central residents. Related to this, Garreau (1992) documents the rise of 'edge cities' in late 20th century US. More recently, Couture and Handbury (2020) and Couture *et al.* (2019) document a resurgence of central cities in the US.

Inspection of Figure 2 shows that this basic pattern occurs along any possible equilibrium path as returns to scale increases: (1) land rents are nearly the same in the center and the periphery for $\gamma$ small; (2) as $\gamma$ increases central land rents enter a region of rapid or discontinuous increase corresponding the industrial revolution; (3) as $\gamma$ increases past the region of rapid or discontinuous increase, as in the late 20th century, land rents in the center and periphery gradually equalize as employment becomes more suburban. Note that in this region of $\gamma$ multiple equilibria are possible, and so this is a region where we would expect to see more complex cities, again, like what emerged in the late 20th century.

Summing up, while both our model and our description of history are stylized, important features of the history of cities can be rationalized by our model and a trend upward in the strength of returns to scale.

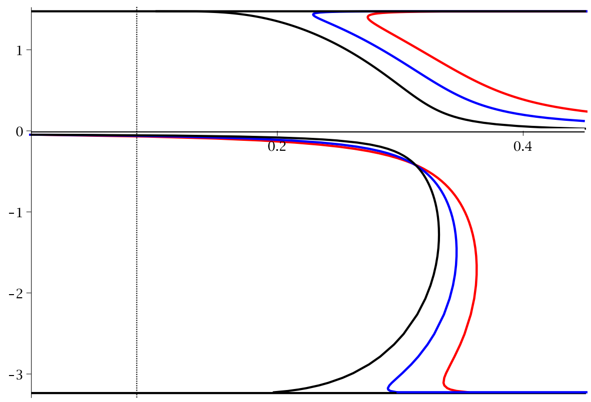
# 7 Spillovers

In the model presented in Section 2 we restrict increasing returns to scale to operate only within a location. Yet, it is common in the urban economics and QSM literatures to model local productivity as a distance discounted sum of economic activity. We here investigate the implications of such spatial externalities.

To accomplish this, we make a single change to the model described in Section 2. In place of the definition of local productivity given by (12) we assume that productivity at $j = 0,1$ is given by:

$$A_0 = (L_0 + 2\delta L_1)^\gamma, \qquad A_1 = (\delta L_0 + (1+\delta^2)L_1)^\gamma. \tag{35}$$

Figure 5: Equilibrium with productivity spillovers.




*Notes: Equilibrium correspondences between* $\ln(\rho)^{\gamma}$ *and* $\gamma$*. Black line: reproduced from figure 2a, but also showing corner equilibria. Blue line: like figure 2a, but with* $\delta = 0.02$*. Red line: like figure 2a, but with* $\delta = 0.04$*.*

Here, $\delta \in [0,1)$ describes the rate at which spillovers decay with distance, and $\gamma \geq 0$ continues to describe the rate at which productivity increases with employment in a location. Note that (35) converges to (12) as $\delta \to 0$.

Recalling that the function $f$ given in Proposition 2 does not depend on $\gamma$, it should be unsurprising that this function remains unchanged from (27) under the more general technology. The function $g$, however, does not remain the same. In particular, following logic similar to that used to derive (28), one can derive the analogous equation when spillovers operate:

$$\omega = \rho^{\frac{1}{a}} \left[ \frac{\omega(\rho + 2\phi) + 2\delta(\phi\rho + 1 + \phi^2)}{(1+\delta^2)(\phi\rho + 1 + \phi^2) + \delta\omega(\rho + 2\phi)} \right]^{\frac{\gamma\varepsilon}{\alpha}}. \tag{36}$$

We have focussed our efforts on numerical solutions to (36) and (27) for small values of $\delta$. This is the empirically relevant range for $\delta$ and provides a basis for thinking about whether equilibrium responds continuously to the introduction of small cross location productivity spillovers.

Figure 5 illustrates typical results. This figure is based on the same parameters as Figure 2a ($\alpha = 0.6$; $\beta = 0.3$; $\varepsilon = 7$; $\tau = 1.1$). The heavy black line reproduces equilibrium correspondence between $\rho$ and $\gamma$ from Figure 2a, where spillovers are zero. The blue and red lines are identical, but assumes $\delta = 0.02$ and 0.04, respectively. Figure 5 suggests two conclusions. First, that the corner equilibria are not robust to the inclusion of a small spillover. Rather, with small spillovers, these equilibria become 'near corner' interior equilibria. Second, $\rho^*$ at a corner equilibrium responds continuously to the introduction of a small spatial productivity spillover.

Summing up, numerical analysis of the economy with a small amount of spatial productivity spillovers suggests that equilibrium responds continuously to the introduction of such spillovers. In other words, we should not expect qualitatively different equilibrium behavior to arise as a consequence of low levels of productivity spillovers, particularly at the interior equilibria.

## 8 Conclusion

Understanding how people arrange themselves when they are free to choose work and residence locations, when commuting is costly, and when increasing returns to scale affect production, is one of the defining problems of urban economics. We address this problem by combining the discrete choice models employed by the recent quantitative spatial models literature, with the stylized geographies of classical urban economics. This permits a complete description of equilibria, throughout the parameter space.

Equilibrium behavior is surprising and interesting for a number of reasons. First, comparative statics as returns to scale increase contradict the conventional wisdom: increasing returns to scale in production can cause dispersion as well as agglomeration. Second, comparative statics on commuting costs also contradict conventional wisdom. As in the monocentric model, reductions in commuting costs can lead to dispersed economic activity, but they can also lead to greater concentration. Third, households' heterogeneity does not affect the city in a monotone way.

By combining classical urban economics and the QSM toolbox, we have also learned about how the two frameworks differ. Introducing agents with heterogenous preferences over work-residence location pairs results in an average preference for central work and residence. These average preferences lead to centralized employment and residence even in the absence of increasing returns to scale. Neither preference is present in the older literature.

Our results also seem to have implications for quantitative exercises. First, we provide a basis for thinking about the extent to which quantitative comparative statics primarily reflect features of a particular data set, or if they are a direct consequence of the interaction of commute costs, returns to scale, and preference heterogeneity in the model. For example, in the absence of returns to scale our model requires lower central than peripheral wages because wages capitalize the average preference for central employment. To match the modern day empirical regularity of higher central than peripheral wages, we require that returns to scale be high enough to more than offset this capitalization. It follows that the returns to scale parameter in this model does not correspond to the elasticity we obtain from regressing wages on density, a common empirical measure of scale economies.

Second, our back of the envelope calculation suggests that quantitative exercises may well be within estimation error of the weak-moderate returns to scale threshold. This is the region of the parameter space where the model's behavior is most complicated. Inspection of Figure 2 indicates that (depending on which regime the economy is in) that catastrophic discontinuities, instability, and stable continuity are all consistent with equilibrium around this threshold. This suggests that at least a rudimentary exploration of what happens when the economy crosses from the weak to moderate increasing returns to scale region is of interest.

Third, one of the more common uses of quantitative models is to evaluate comparative statics for commuting costs. Our results make clear that these comparative statics are complicated and may be sensitive to changes in the values of structural parameters. Our simple model does not provide a basis for thinking about how sensitive richer empirical models may be to this problem, but it does suggest that the possibility should at least be considered. This suggests the importance of robustness checks in which the values of these parameters are permitted to vary.

Last, we show that the qualitative behavior of our model as returns to scale increase can reproduce many of the qualitative features observed over the last 500 years of urban history in the Western world. Urbanization and increasing productivity are surely two of the most important economic phenomena in history and they appear to have been closely linked.

## References


Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6):2127–2189.

Allen, R. C. (2001). The great divergence in european wages and prices from the middle ages to the first world war. *Explorations in Economic History*, 38(4):411–447.

Allen, T. and Arkolakis, C. (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics*, 129(3):1085–1140.

Allen, T., Arkolakis, C., and Li, X. (2015). Optimal city structure. *Yale University, mimeograph*.

Allen, T., Arkolakis, C., and Li, X. (2020a). On the equilibrium properties of network models with heterogeneous agents. Technical report, National Bureau of Economic Research.

Allen, T., Arkolakis, C., and Takahashi, Y. (2020b). Universal gravity. *Journal of Political Economy*, 128(2):393–433.

Bairoch, P. (1988). *Cities and economic development: from the dawn of history to the present*. University of Chicago Press.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics*, 122(2):775–805.

Cantoni, D. and Yuchtman, N. (2014). Medieval universities, legal institutions, and the commercial revolution. *The Quarterly Journal of Economics*, 129(2):823–887.

Clark, C. (1951). Urban population densities. *Journal of the Royal Statistical Society*, 114(4):490–496.

Costa, D. L. (2015). Health and the economy in the united states from 1750 to the present. *Journal of Economic Literature*, 53(3):503–70.

Couture, V., Gaubert, C., Handbury, J., and Hurst, E. (2019). Income growth and the distributional effects of urban spatial sorting. Technical report, National Bureau of Economic Research.

Couture, V. and Handbury, J. (2020). Urban revival in america. *Journal of Urban Economics*, 119:103267.

de la Croix, D., Doepke, M., and Mokyr, J. (2018). Clans, guilds, and markets: Apprenticeship institutions and growth in the preindustrial economy. *The Quarterly Journal of Economics*, 133(1):1–70.

de Vries, J. (1984). *European Urbanization, 1500-1800*. Routledge.

Dittmar, J. E. (2011). Information technology and economic change: the impact of the printing press. *The Quarterly Journal of Economics*, 126(3):1133–1172.

Fujita, M. (1989). *Urban economic theory: Land use and city size*. Cambridge University Press.

Garreau, J. (1992). *Edge city: Life on the new frontier*. Anchor.

Glaeser, E. L. and Kahn, M. E. (2004). Sprawl and urban growth. In *Handbook of Regional and Urban Economics*, volume 4, pages 2481–2527. Elsevier.

Heblich, S., Redding, S. J., and Sturm, D. M. (2020). The making of the modern metropolis: evidence from london. *The Quarterly Journal of Economics*, 135(4):2059–2133.

Lucas, R. E. and Rossi-Hansberg, E. (2002). On the internal structure of cities. *Econometrica*, 70(4):1445–1476.

Monte, F., Redding, S. J., and Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. *American Economic Review*, 108(12):3855–90.

Ogawa, H. and Fujita, M. (1980). Equilibrium land use patterns in a nonmonocentric city. *Journal of Regional Science*, 20(4):455–475.

Redding, S. J. and Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, 9:21–58.

Redding, S. J. and Turner, M. A. (2015). Transportation costs and the spatial organization of economic activity. *Handbook of regional and urban economics*, 5:1339–1398.

Starrett, D. (1978). Market allocations of location choice in a model with free mobility. *Journal of economic theory*, 17(1):21–37.

# Appendix

### *A. Proof of Proposition 1*

We first evaluate $\kappa$ for our three location city. Using (5), symmetry and (19) we have that

$$\kappa = \left[ R_1^{\beta\varepsilon} W_1^{-\varepsilon} \left( 2(1+\phi^2) + 2\phi\omega + 2\phi\rho + \omega\rho \right) \right]^{-1}. \tag{A.1}$$

Using the expression for utility maximizing choice shares (20, the definition of residential population (7), along with symmetry and (19), we have:

$$M_0 = s_{00} + 2s_{01} = \kappa R_1^{\beta\varepsilon} W_1^{-\varepsilon} (\rho(2\phi + \omega)).$$

Substituting for $\kappa$ from (A.1), we have $M_0$ as a function of $\rho$ and $\omega$. Under symmetry, the residential pattern satisfies $2M_1 + M_0 = 1$. Substituting from (23) gives $M_1$ as a function of $\rho$ and $\omega$.

Using the expression for utility maximizing choice shares (20), the definition of employment (8), along with symmetry and (19), we obtain:

$$L_0 = s_{00} + 2s_{10} = \kappa R_1^{\beta\varepsilon} W_1^{-\varepsilon} (\rightarrow (2\phi + \rho)). \tag{A.2}$$

Substituting for $\kappa$ from (A.1), we have $L_0$ as a function of $\rho$ and $\omega$. Under symmetry, the employment pattern satisfies $2L_1 + L_0 = 1$. Substituting from (A.2) gives $L_1$.

To evaluate expressions for residential and commercial land, evaluate (10) for our three location model to get

$$R_0 H_0 = \beta\kappa \left( W_0^{1+\varepsilon} R_0^{-\beta\varepsilon} + 2\phi W_1^{1+\varepsilon} R_0^{-\beta\varepsilon} \right),$$

$$R_1 H_1 = \beta\kappa \left[ \phi W_0^{1+\varepsilon} R_1^{-\beta\varepsilon} + (1+\phi^2) W_1^{1+\varepsilon} R_1^{-\beta\varepsilon} \right].$$

Substituting from (A.1) and using (19) we have that

$$H_0 = \beta \frac{W_0}{R_0} \frac{\rho\omega + 2\phi\rho\omega^{-\frac{1}{\varepsilon}}}{\omega\rho + 2\phi\rho + 2\phi\omega + 2(1+\phi^2)},$$

$$H_1 = \beta \frac{W_1}{R_1} \frac{\phi\omega^{\frac{1+\varepsilon}{\varepsilon}} + (1+\phi^2)}{\omega\rho + 2\phi\rho + 2\phi\omega + 2(1+\phi^2)}.$$

Using the Cobb-Douglas property (15) and the land market clearing conditions (9), we get

$$1 - N_0 = a\frac{N_0}{L_0}\frac{\rho\omega + 2\phi\rho\omega^{-\frac{1}{\varepsilon}}}{\omega\rho + 2\phi\rho + 2\phi\omega + 2(1+\phi^2)},$$

$$1 - N_1 = a\frac{N_1}{L_1}\frac{\phi\omega^{\frac{1+\varepsilon}{\varepsilon}} + (1+\phi^2)}{\omega\rho + 2\phi\rho + 2\phi\omega + 2(1+\phi^2)}.$$

Solving for $N_0$ and $N_1$ and using the expressions $L_0$ and $L_1$ from (24), we arrive at the commercial land use pattern,

$$N_0 = \frac{\rho + 2\phi}{(1+a)\rho + 2\phi + 2a\phi\rho\omega^{-\frac{1+\varepsilon}{\varepsilon}}},$$

$$N_1 = \frac{\phi\rho + (1+\phi^2)}{\phi\rho + a\phi\omega^{\frac{1+\varepsilon}{\varepsilon}} + (1+a)(1+\phi^2)}.$$

Substituting these expressions into the land market clearing conditions, $H_i + N_i = 1$, we find the housing pattern (25) and (26). Q.E.D.

***B. Proof of Proposition 2***

The land rent equilibrium condition (11) at $i = 0{,}1$ leads to

$$R_0 = \left[\frac{\beta\kappa}{H_0}\left(W_0^{1+\varepsilon} + 2\phi W_1^{1+\varepsilon}\right)\right]^{\frac{1}{1+\beta\varepsilon}}, \tag{B.1}$$

$$R_1 = \left[\frac{\beta\kappa}{H_1}\left(\phi W_0^{1+\varepsilon} + (1+\phi^2) W_1^{1+\varepsilon}\right)\right]^{\frac{1}{1+\beta\varepsilon}}. \tag{B.2}$$

Multiplying by $R_j$ both sides of the land market balance condition (9), we get:

$$R_0H_0 + R_0N_0 = R_0H_0 + \frac{1-\alpha}{\alpha}W_0L_0 = R_0, \tag{B.3}$$

$$R_1H_1 + R_1N_1 = R_1H_1 + \frac{1-\alpha}{\alpha}W_1L_1 = R_1. \tag{B.4}$$

Dividing (B.3) over (B.4), we obtain:

$$r = \frac{R_0}{R_1} = \frac{R_0H_0 + \frac{1-\alpha}{\alpha}W_0L_0}{R_1H_1 + \frac{1-\alpha}{\alpha}W_1L_1}. \tag{B.5}$$

It then follows from (B.1) – (B.2) that

$$R_0H_0 = \beta\kappa\left(W_0^{1+\varepsilon}R_0^{-\beta\varepsilon} + 2\phi W_1^{1+\varepsilon}R_0^{-\beta\varepsilon}\right), \tag{B.6}$$

$$R_1H_1 = \beta\kappa\left[\phi W_0^{1+\varepsilon}R_1^{-\beta\varepsilon} + (1+\phi^2)W_1^{1+\varepsilon}R_1^{-\beta\varepsilon}\right]. \tag{B.7}$$

Using (20), the labor market balance conditions at $i = 0{,}1$ are given by

$$L_0 = s_{00} + 2s_{10} = \kappa W_0^{\varepsilon}\left(R_0^{-\beta\varepsilon} + 2\phi R_1^{-\beta\varepsilon}\right),$$

$$L_1 = s_{11} + s_{01} + s_{-11} = \kappa W_1^{\varepsilon}\left[(1+\phi^2)R_1^{-\beta\varepsilon} + \phi R_0^{-\beta\varepsilon}\right],$$

so that

$$W_0L_0 = \kappa\left(W_0^{1+\varepsilon}R_0^{-\beta\varepsilon} + 2\phi W_0^{1+\varepsilon}R_1^{-\beta\varepsilon}\right), \tag{B.8}$$

$$W_1L_1 = \kappa\left[\phi W_1^{1+\varepsilon}R_0^{-\beta\varepsilon} + (1+\phi^2)W_1^{1+\varepsilon}R_1^{-\beta\varepsilon}\right]. \tag{B.9}$$

Plugging into (B.5) the expressions for $R_jH_j$ ($j = 0{,}1$) given by (B.6)-(B.7) and the expressions for $W_jL_j$ ( $j = 0{,}1$) given by (B.8)-(B.9), we get after simplifications:

$$r = \frac{(1+a)w^{1+\varepsilon}r^{-\beta\varepsilon} + 2a\phi r^{-\beta\varepsilon} + 2\phi w^{1+\varepsilon}}{a\phi w^{1+\varepsilon} + \phi r^{-\beta\varepsilon} + (1+a)(1+\phi^2)}. \tag{B.10}$$

Combining (12) with (13)-(14), we get:

$$w = \ell^{\gamma}\left(\frac{n}{\ell}\right)^{1-\alpha} \qquad \text{and} \qquad r = \ell^{\gamma}\left(\frac{n}{\ell}\right)^{-\alpha}. \tag{B.11}$$

Dividing (B.8) by (B.9) yields:

$$\ell = w^{\varepsilon}\frac{r^{-\beta\varepsilon} + 2\phi}{\phi r^{-\beta\varepsilon} + (1+\phi^2)}. \tag{B.12}$$

Using (B.11) and (B.12), we get:

$$w^{\alpha}r^{1-\alpha} = \ell^{\gamma} = \left[w^{\varepsilon}\frac{r^{-\beta\varepsilon} + 2\phi}{\phi r^{-\beta\varepsilon} + (1+\phi^2)}\right]^{\gamma},$$

or, equivalently,

$$w = r^{-\frac{1-\alpha}{\alpha}}\left[w^{\varepsilon}\frac{r^{-\beta\varepsilon} + 2\phi}{\phi r^{-\beta\varepsilon} + (1+\phi^2)}\right]^{\frac{\gamma}{\alpha}}. \tag{B.13}$$

Since $\omega > 0$, the conditions (B.10) and (B.13) can be reformulated in terms of $(\rho,\omega)$:

$$\rho = \left[\frac{(1+a)\rho\omega^{\frac{1+\varepsilon}{\varepsilon}} + 2a\phi\rho + 2\phi\omega^{\frac{1+\varepsilon}{\varepsilon}}}{\phi\rho + a\phi\omega^{\frac{1+\varepsilon}{\varepsilon}} + (1+\phi^2)(1+a)}\right]^{-\beta\varepsilon}, \tag{B.14}$$

$$\omega = \rho^{\frac{1}{a}}\left(\omega\frac{\rho + 2\phi}{\phi\rho + 1 + \phi^2}\right)^{\frac{\gamma\varepsilon}{\alpha}}. \tag{B.15}$$

Solving (B.14) and (B.15) for $\omega^{\frac{1+\varepsilon}{\varepsilon}}$ yields (27) and (28). Q.E.D.

## *C. Proof of Lemma 1*

(i) The denominator of $f(\rho)$ is an increasing function of $\rho$, which is unbounded above and negative at $\rho = 0$. Therefore, the equation

$$D(\rho) \equiv (1+a)\rho^{1+\frac{1}{\beta\varepsilon}} + 2\phi\rho^{\frac{1}{\beta\varepsilon}} - a\phi = 0$$

has a unique positive root $\rho_0$. Clearly, $D(\rho) > 0$ if and only if $\rho > \rho_0$. Since $D(1) > 0$, it must be that $\rho_0 < 1$.

(ii) It is readily verified using (27) that the numerator of $f(\rho)$ is a concave function which is positive at zero. It increases in the vicinity of 0 and then starts decreasing and goes to $-\infty$ as $\rho \to \infty$. Hence, the equation

$$N(\rho) \equiv \phi\rho - 2a\phi\rho^{1+\frac{1}{\beta\varepsilon}} + (1+\phi^2)(1+a) = 0$$

has a unique positive root $\rho_1$. Clearly, $N(\rho) < 0$ if and only if $\rho > \rho_1$. Since $N(1) > 0$, it must be that $\rho_1 > 1$.

(iii) We have $f(\rho) > 0$ if and only if only if the numerator $N(\rho)$ and the denominator $D(\rho)$ have the same sign. This holds if and only if $\rho_0 < \rho < \rho_1$. The interval $(\rho_0, \rho_1)$ is not empty because $\rho_0 < 1 < \rho_1$. As a result, $f(\rho) > 0$ if and only if $\rho \in (\rho_0, \rho_1)$.

(iv) It remains to show that $f(\rho)$ decreases from $\infty$ to 0 over $(\rho_0, \rho_1)$.

**Step 1.** We first show that

$$\psi(x,y) \equiv \frac{(1+a)xy + 2a\phi x + 2\phi y}{\phi x + a\phi y + (1+a)(1+\phi^2)} \tag{C.1}$$

is increasing in both $x$ and $y$ for all $(x,y) \in R^2_{++}$. We have

$$\frac{\partial\psi(x,y)}{\partial x} = \frac{(1+a)a\phi y^2 + \left[(1+a)^2(1+\phi^2) + 2a^2\phi^2 - 2\phi^2\right] y + 2a\phi(1+a)(1+\phi^2)}{\left[\phi x + a\phi y + (1+a)(1+\phi^2)\right]^2} > 0,$$

$$\frac{\partial\psi(x,y)}{\partial y} = \frac{(1+a)a\phi x^2 + \left[(1+a)^2(1+\phi^2) - 2a^2\phi^2 + 2\phi^2\right] x + 2a\phi(1+a)(1+\phi^2)}{\left[\phi x + a\phi y + (1+a)(1+\phi^2)\right]^2} > 0,$$

because, as $0 < \phi < 1$,

$$(1+a)^2(1+\phi^2) + 2a^2\phi^2 - 2\phi^2 > 1 + \phi^2 - 2\phi^2 = 1 - \phi^2 > 0,$$

$$(1+a)^2(1+\phi^2) - 2a^2\phi^2 + 2\phi^2 > a^2(1+\phi^2) - 2a^2\phi^2 = a^2(1-\phi^2) > 0.$$

**Step 2.** It is readily verified that $f(\rho)$ satisfies the following identity:

$$\rho = \left[\frac{(1+a)\rho f(\rho) + 2a\phi\rho + 2\phi f(\rho)}{\phi\rho + a\phi f(\rho) + (1+\phi^2)(1+a)}\right]^{-\beta\varepsilon} = [\psi(\rho, f(\rho))]^{-\beta\varepsilon},$$

where $\psi$ is defined by (C.1). When $\rho$ increases, the RHS of this expression also increases, which means that $\psi(\rho, f(\rho))$ decreases with $\rho$. Since $\psi$ increases with $\rho$ and $f(\rho)$, this is possible only if $f(\rho)$ is decreasing. Q.E.D.

### *D. Proof of Lemma 2*

Part (i) follows from combining (28) with $0 < \rho_0 < \rho_1 < \infty$. Parts (ii) and (iii) are obtained by differentiating $g$ with respect to $\rho$. Part (iv) holds because $g(\rho;\gamma)$ may be rewritten as follows:

$$g(\rho;\gamma) = \left[\left(\rho^{\frac{1}{a}} \frac{\rho + 2\phi}{1 + \phi\rho + \phi^2}\right)^{\frac{\gamma_m}{\gamma_m - \gamma}} \cdot \left(\frac{\rho + 2\phi}{1 + \phi\rho + \phi^2}\right)^{-1}\right]^{\frac{1+\varepsilon}{\varepsilon}}. \tag{D.1}$$

Q.E.D.

### *E. Proof of Lemma 3*

It follows from the proof of Lemma 1 that $\rho_0$ is the unique solution of

$$D(\rho) \equiv (1+a)\rho^{\frac{1+\beta\varepsilon}{\beta\varepsilon}} + 2\phi\rho^{\frac{1}{\beta\varepsilon}} - a\phi = 0. \tag{E.1}$$

The expressions (30) and (E.1) imply that $\rho_L$ and $\rho_0$ are functions of $a$. We next show that $\rho_0$ and $\rho_L$ vary with $a$ as follows,

$$\lim_{a\to 0} \rho_L = 1, \quad \frac{d\rho_L}{da} < 0, \quad \lim_{a\to\infty} \rho_L = 1 - \phi,$$

$$\lim_{a\to 0} \rho_0 = 0, \quad \frac{d\rho_0}{da} > 0, \quad \lim_{a\to\infty} \rho_0 = \phi^{\frac{\beta\varepsilon}{1+\beta\varepsilon}}.$$

We can show that $\rho_0$ (resp., $\rho_L$) increases (resp., decreases) in $a$ by applying the implicit function theorem to $D(\rho) = 0$ (resp., (30)). Observe further that, when $a \to \infty$ (resp., $a \to 0$), dividing $D(\rho) = 0$ by $a$ and taking the limit yields $\rho_0 = \phi^{\beta\varepsilon/(1+\beta\varepsilon)}$ (resp., $\rho_0 = 1$). Last, when $a \to \infty$ (resp., $a \to 0$), taking (E.1) at the power $a$ and the limit yields $\rho_L = 1 - \phi$ (resp., $\rho_L = 1$).

To determine where $\rho_0$ and $\rho_L$ intersect, we compare $\lim_{a\to\infty} \rho_L$ and $\lim_{a\to\infty} \rho_0$ by considering the equation

$$\phi^{\beta\varepsilon/(1+\beta\varepsilon)} + \phi = 1. \tag{E.2}$$

Differentiating the LHS of (E.2) with respect to $\phi$ shows that it increases from 0 to 2 when $\phi$ increases from 0 to 1. The intermediate value theorem then implies that, for any given $\beta\varepsilon$, (E.2) has a unique solution $\overline{\phi}(\beta\varepsilon) \in (0,1)$, which increases with $\beta\varepsilon$.

The inequality $\rho_0 \leq \rho_L$ holds if $\phi^{\beta\varepsilon/(1+\beta\varepsilon)} \leq 1 - \phi$, which amounts to $\phi \leq \overline{\phi}(\beta\varepsilon)$. If $\overline{\phi} < \phi \leq 1$, then there exists a unique value $\overline{a} > 0$ that solves the condition $\rho_L(a) = \rho_0(a)$. Consequently, if $a < \overline{a}$, then $\rho_0 \leq \rho_L$. If $a \geq \overline{a}$, then $\rho_0 > \rho_L$. Summing up, $\rho_0 \leq \rho_L$ if $\phi \leq \overline{\phi}$ or $a \leq \overline{a}$, and $\rho_0 > \rho_L$ when both conditions fail. Q.E.D.

***F. Proof of Proposition 3***

Using the properties of $f$ and $g$ given in Lemmas 1 and 2, the intermediate value theorem implies the existence of an equilibrium when $\gamma \neq \gamma_m$. The case where $\gamma = \gamma_m$ is discussed in Section 6.1.2.

Furthermore, there exist two corner equilibria for all $\gamma > 0$: (i) $\omega^* \to \infty$ and $\rho^* = \rho_1$, and thus employment is evenly distributed between peripheries, and (ii) $\omega^* = 0$ and $\rho^* = \rho_0$, so that employment is fully concentrated at the center $i = 0$. Only two employment patterns may be symmetric corner equilibria: (0,1,0) and (1/2,0,1/2).

**Case 1**: $\mathbf{L}^* = (1/2,0,1/2)$. The demand for commercial land at $j = 0$ is zero ($N_0^* = 0$). Hence, the land rent equilibrium condition (11) at $i = 0$ and $i = 1$ amount to

$$R_0 = \left[\frac{\beta\kappa}{H_0} 2\phi W_1^{1+\varepsilon}\right]^{\frac{1}{1+\beta\varepsilon}} > 0, \tag{F.1}$$

$$R_1 = \left[\frac{\beta\kappa}{H_1} (1+\phi^2) W_1^{1+\varepsilon}\right]^{\frac{1}{1+\beta\varepsilon}} > 0. \tag{F.2}$$

Multiplying both sides of the land balance conditions (F.1) and (F.2) by, respectively, $R_0$ and $R_1$, we get:

$$R_0 H_0 = R_0, \tag{F.3}$$

$$R_1 H_1 + R_1 N_1 = R_1 H_1 + \frac{1-\alpha}{\alpha} W_1 L_1 = R_1. \tag{F.4}$$

Dividing (F.3) over (F.4), we obtain:

$$r = \frac{R_0}{R_1} = \frac{R_0}{R_1 H_1 + \frac{1-\alpha}{\alpha} W_1 L_1}. \tag{F.5}$$

It then follows from (F.1)-(F.2) that

$$R_0 = \beta\kappa 2\phi W_1^{1+\varepsilon} R_0^{-\beta\varepsilon}, \tag{F.6}$$

$$R_1 H_1 = \beta\kappa(1+\phi^2) W_1^{1+\varepsilon} R_1^{-\beta\varepsilon}. \tag{F.7}$$

The labor market balance condition at $j = 1$ is given by

$$L_1 = s_{11} + s_{01} + s_{-11} = \kappa W_1^{\varepsilon}\left[(1+\phi^2)R_1^{-\beta\varepsilon} + \phi R_0^{-\beta\varepsilon}\right],$$

so that

$$W_1 L_1 = \kappa\left[\phi W_1^{1+\varepsilon} R_0^{-\beta\varepsilon} + (1+\phi^2) W_1^{1+\varepsilon} R_1^{-\beta\varepsilon}\right]. \tag{F.8}$$

Plugging (F.6)-(F.7) and (F.8) into (F.5), we get after simplifications:

$$r = \frac{2a\phi r^{-\beta\varepsilon}}{\phi r^{-\beta\varepsilon} + (1+a)(1+\phi^2)},$$

or, equivalently,

$$\rho = \left[\frac{2a\phi\rho}{\phi\rho + (1+a)(1+\phi^2)}\right]^{-\beta\varepsilon}$$

whose unique solution is $\rho^* = \rho_1$. In other words, employment is evenly concentrated in the two peripheries and the central location is specialized in residential activities.

**Case 2**: $\mathbf{L}^* = (0,1,0)$. Following the same argument as in Case 1, we end up with the equation:

$$\rho = \left[\frac{(1+a)\rho + 2\phi}{a\phi}\right]^{-\beta\varepsilon}$$

whose unique solution is given by $\rho^* = \rho_0$. Therefore, $\rho^* = \rho_0$ is an equilibrium if and only if $L_0^* = 1$.

Finally, we show that $M_i^* > 0$ for all $i$. Assume that $R_i^* = 0$ at $i$. Since there is a location $j \in \mathcal{I}$ such that $W_j^* > 0$, workers who choose the pair $ij$ enjoys an infinite utility level, which implies $s_{ij} > 0$. These workers' land demand is thus infinite while the land supply is finite, a contradiction. Q.E.D.

### *G. Proof of Proposition 4*

Proposition 3 implies that there is no corner equilibria when $\gamma = 0$. Since $g(\rho; 0)$ increases from 0 to 1 as $\rho$ increases over $[0,1]$, Lemma 1 implies that the two curves must cross exactly once. Furthermore, the intersection must occur strictly between $\rho_0$ and 1, which implies $0 < \rho^* < 1$. Since $g < 1$ over this interval, it must be that $\omega^* < 1$, while (33) implies $\omega^* > 0$.

Finally, we show that (32) is sufficient for the equilibrium employment pattern to be bell-shaped. Since $0 < \rho^* < 1$, we may assume throughout that $\rho \in (0,1)$. Note that the equilibrium employment pattern is bell-shaped if and only if $\ell^* > 1$, while (32) is equivalent to $b > 1$.

Some tedious calculations show that the equilibrium condition $f(\rho) = g(\rho; 0)$ may be rewritten as follows:

$$\frac{\left(\frac{a}{1+a}\rho^{-b} + \frac{1}{1+a}\rho^{-1}\right)^{-1} + 2\phi}{\phi\left(\frac{a}{1+a}\rho^{b} + \frac{1}{1+a}\rho\right) + 1 + \phi^2}\left(\frac{a}{1+a}\rho^{1+\frac{1}{\beta\varepsilon}} + \frac{1}{1+a}\rho^{b+\frac{1}{\beta\varepsilon}}\right) = 1. \tag{G.1}$$

Since $1/x$ is convex, for every $\rho < 1$ Jensen's inequality implies

$$\left(\frac{a}{1+a}\rho^{-b} + \frac{1}{1+a}\rho^{-1}\right)^{-1} < \frac{a}{1+a}\rho^{b} + \frac{1}{1+a}\rho < \rho. \tag{G.2}$$

Plugging (G.2) into (G.1) leads to

$$1 < \frac{\frac{a}{1+a}\rho^{b} + \frac{1}{1+a}\rho + 2\phi}{\phi\left(\frac{a}{1+a}\rho^{b} + \frac{1}{1+a}\rho\right) + 1 + \phi^2}\left(\frac{a}{1+a}\rho^{1+\frac{1}{\beta\varepsilon}} + \frac{1}{1+a}\rho^{b+\frac{1}{\beta\varepsilon}}\right).$$

Using $b > 1$ yields

$$\frac{a}{1+a}\rho^{b} + \frac{1}{1+a}\rho < \frac{a}{1+a}\rho + \frac{1}{1+a}\rho = \rho.$$

Since the function $\frac{x+2\phi}{\phi x+1+\phi^2}$ is increasing for all $x \geq 0$, we obtain

$$1 < \frac{\rho + 2\phi}{\phi\rho + 1 + \phi^2}\left(\frac{a}{1+a}\rho^{1+\frac{1}{\beta\varepsilon}} + \frac{1}{1+a}\rho^{b+\frac{1}{\beta\varepsilon}}\right). \tag{G.3}$$

As (32) implies

$$\frac{1}{a} < 1 + \frac{1}{\beta\varepsilon} < b + \frac{1}{\beta\varepsilon},$$

while $\rho^* < 1$, we have

$$\frac{a}{1+a}\left(\rho^*\right)^{1+\frac{1}{\beta\varepsilon}} + \frac{1}{1+a}\left(\rho^*\right)^{b+\frac{1}{\beta\varepsilon}} < \left(\rho^*\right)^{\frac{1}{a}}.$$

Replacing the bracketed term in (G.3), we obtain the inequality:

$$1 < \left(\rho^*\right)^{\frac{1}{a}}\frac{\rho^* + 2\phi}{\phi\rho^* + 1 + \phi^2},$$

which is equivalent to $\rho^* > \rho_L$, hence $\ell^* > 1$ (see (31)). Q.E.D.

## *H. Proof of Proposition 5*

Under weak IRS, given Lemmas 1 and 2, $f$ and $g$ must intersect exactly once. Furthermore, because $f(1) < 1 < g(1;\gamma)$, the intersection must occur at $\rho^* < 1$. Since $\rho^* > \rho_L$, we have

$$f(\rho_L) > f(\rho^*) = g(\rho^*;\gamma) > g(\rho_L) \tag{H.1}$$

because $f$ is decreasing by Lemma 1 and $g$ is increasing in $\rho$ by Lemma 2. As shown by (D.1), $g(\rho_L)$ is independent of $\gamma$. Combining this with (H.1), we obtain $f(\rho_L) - g(\rho_L;\gamma) > 0$. Since $f(\rho^*) - g(\rho^*;\gamma) = 0$ while $f - g$ is decreasing by Lemmas 1 and 2, we have $\rho_L < \rho^*$ for all $\gamma < \alpha/\varepsilon$, which amounts to $\ell^* > 1$.

We now study the impact of $\gamma$ on (i) $\rho^*$, (ii) $\omega^*$and (iii) $\ell^*$.

(i) Since $\partial g(\rho;\gamma)/\partial\gamma > 0$, applying the implicit function theorem to (29) leads to

$$\frac{d\rho^*}{d\gamma} = \left.\frac{\partial g(\rho;\gamma)/\partial\gamma}{\partial f'(\rho)/\partial\rho - \partial g(\rho;\gamma)/\partial\rho}\right|_{\rho=\rho^*} < 0,$$

where the numerator is positive because $\rho^* > \rho_L$ while the denominator is negative because $f(\rho)$ is decreasing and $g(\rho;\gamma)$ is increasing in $\rho$.

(ii) Differentiating (27) with respect to $\gamma$, we obtain:

$$\frac{1+\varepsilon}{\varepsilon}\omega^{\frac{1}{\varepsilon}}\frac{d\omega^*}{d\gamma} = \frac{df}{d\rho}\frac{d\rho^*}{d\gamma} > 0.$$

(iii) Observe that, combining (31) with (D.1), the equilibrium condition (29) can be restated as

$$\ell^{\frac{1+\varepsilon}{\varepsilon}} = \left(\frac{\rho+2\phi}{\phi\rho+1+\phi^2}\right)^{\frac{1+\varepsilon}{\varepsilon}} f(\rho). \tag{H.2}$$

Since $f(\rho)$ can be decomposed as

$$f(\rho) = \rho^{-\frac{1}{\beta\varepsilon}} \cdot \frac{\phi\rho+1+\phi^2}{\rho+2\phi} \cdot \frac{1+a\frac{1+\phi^2-2\phi\rho^{1+\frac{1}{\beta\varepsilon}}}{\phi\rho+1+\phi^2}}{1+a\frac{\rho-\phi\rho^{-\frac{1}{\beta\varepsilon}}}{\rho+2\phi}}, \tag{H.3}$$

Plugging (H.3) into (H.2) leads to

$$\ell^{\frac{1+\varepsilon}{\varepsilon}} = \left(\frac{\rho^{-\frac{1-\beta}{\beta}}+2\phi\rho^{-\frac{1}{\beta}}}{\phi\rho+1+\phi^2}\right)^{\frac{1}{\varepsilon}} \cdot \frac{1+a\frac{1+\phi^2-2\phi\rho^{1+\frac{1}{\beta\varepsilon}}}{\phi\rho+1+\phi^2}}{1+a\frac{\rho-\phi\rho^{-\frac{1}{\beta\varepsilon}}}{\rho+2\phi}}.$$

The first term in the RHS clearly decreases in $\rho$. Since the numerator (resp., denominator) of the second term is decreasing (resp., increasing), the second term also decreases in $\rho$. Hence, the RHS is decreasing in $\gamma$. Combining this with $d\rho^*/d\gamma < 0$, we obtain $d\ell^*/d\gamma > 0$. Q.E.D.

***I. Proof of Proposition 6***

(i) Consider first the case when commuting costs are high. It then follows from (30) and Lemma 3 that $\rho_0 < \rho_L < 1 < \rho_1$. Therefore, for $\Delta > 0$ sufficiently small, we have:

$$\rho_0 + \Delta < \rho_L - \Delta < \rho_L + \Delta < 1 < \rho_1.$$

If $\gamma$ is sufficiently close to $\alpha/\varepsilon$ (but still such that $\gamma > \alpha/\varepsilon$ holds), Lemma 2 implies the following inequalities:

$$\begin{aligned}
g(\rho_0 + \Delta; \gamma) &< f(\rho_0 + \Delta), \\
g(\rho_L - \Delta; \gamma) &> f(\rho_L - \Delta), \\
g(\rho_L + \Delta; \gamma) &< f(\rho_L + \Delta), \\
g(\rho_1; \gamma) &> f(\rho_1) = 0,
\end{aligned}$$

where the last inequality holds because (28) implies that, for $\gamma > \alpha/\varepsilon$, $g(\rho; \gamma) > 0$ for all $\rho > 0$ while $f(\rho_1) = 0$ for any $\gamma$ by definition of $\rho_1$. Therefore, by continuity of $f$ and $g$, (29) has at least *three* distinct solutions, which we denote as follows:

$$\rho_1^\star > \rho_2^\star > \rho_3^\star.$$

Furthermore, the properties of function $g$ imply the following:

$$\lim_{\gamma\varepsilon \searrow \alpha} \rho_1^\star = \rho_1,$$

$$\lim_{\gamma\varepsilon \searrow \alpha} \rho_2^\star = \rho_L,$$

$$\lim_{\gamma\varepsilon \searrow \alpha} \rho_3^\star = \rho_0.$$

The solution $\rho_2^\star$ matches the equilibrium of Proposition 5. As for the other two solutions, $\rho^\star$ and $\rho_3^\star$, when $\gamma$ is close enough to $\alpha/\varepsilon$, we have $\rho^\star > 1 > \rho_3^\star$.

As $\gamma \searrow \alpha/\varepsilon$, it follows from Lemma 1 that $f(\rho^\star)$ and $f(\rho_3^\star)$ converge, respectively, to 0 and $\infty$, which implies:

$$\lim_{\gamma\varepsilon \searrow \alpha} \omega_1^\star = 0 \qquad \text{and} \qquad \lim_{\gamma\varepsilon \searrow \alpha} \omega_3^\star = \infty.$$

Hence, $\omega_1^\star < 1 < \omega_3^\star$ when $\gamma\varepsilon$ is close enough to $\alpha$. It then follows from (28) that

$$\lim_{\gamma\varepsilon \searrow \alpha} \ell_1^\star = 0 \qquad \text{and} \qquad \lim_{\gamma\varepsilon \searrow \alpha} \ell_3^\star = \infty.$$

(ii) Consider now the case of high commuting costs. Then, we know from the proof of Proposition 7 that there exists a value $\overline{a} \in (0,1)$ such that

$$\rho_L \leq \rho_0 < 1 < \rho_1 \tag{I.1}$$

is satisfied for $a \geq \overline{a}$, and $\rho_0 < \rho_L < 1 < \rho_1$ holds otherwise. Under (I.1), there is a small $\Delta > 0$ such that the following inequalities hold:

$$g(\rho_1 - \Delta; \gamma) < f(\rho_1 - \Delta),$$
$$g(\rho_1; \gamma) > f(\rho_1) = 0.$$

while $\rho^\star > 1$ when $\gamma$ slightly exceeds $\alpha/\varepsilon$.

Furthermore,

$$\lim_{\gamma\varepsilon \searrow \alpha} (\omega_1^*)^{\frac{\varepsilon}{1+\varepsilon}} = f(\rho_1) = 0.$$

Since $\lim_{\gamma\varepsilon \searrow \alpha} \omega_1^\star = 0$, $\omega_1^\star < 1$ when $\gamma\varepsilon$ is sufficiently close to $\alpha$. Last, using (31), we have:

$$\lim_{\gamma\varepsilon \searrow \alpha} \ell_1^\star = 0.$$

Q.E.D.

### *J. Proof of Proposition 7*

First, we show the existence and uniqueness of an equilibrium. Using (D.1) and (H.3), the equilibrium condition (29) becomes after simplifications:

$$\frac{1}{\phi\rho + 1 + \phi^2} \left( \frac{1 + a\frac{1+\phi^2-2\phi\rho^{1+\frac{1}{\beta\varepsilon}}}{\phi\rho+1+\phi^2}}{1 + a\frac{\rho-\phi\rho^{-\frac{1}{\beta\varepsilon}}}{\rho+2\phi}} \right)^{\lambda} = \rho^{\mu} \frac{\rho}{\rho + 2\phi}, \tag{J.1}$$

where $\lambda$ and $\mu$ are defined by

$$\lambda \equiv \frac{\gamma\varepsilon - \alpha}{\gamma + \alpha} > 0 \qquad \text{and} \qquad \mu \equiv \frac{\gamma\varepsilon - \alpha - (1-\alpha)(1+\varepsilon)}{\beta\varepsilon(\gamma + \alpha)}.$$

The first term of the LHS of (J.1) decreases in $\rho$; the second term also decreases because the numerator decreases while the denominator increases in $\rho$.

Therefore, the LHS of (J.1) is a decreasing function of $\rho$. Furthermore, the RHS of (G.1) increases from 0 to $\infty$ in $\rho$ when $\mu > 0$. It is readily verified that $\mu > 0$ if and only if

$$\gamma > \frac{1+\varepsilon}{(1-\beta)\varepsilon} - \alpha.$$

Hence, (H.1) has a unique solution $\rho^*$.

We now show that $\ell^*$ converges monotonically toward 1 when $\gamma > \gamma_s$ increases. Using (31), we obtain

$$\log \ell^* = -\frac{1}{\gamma\varepsilon/\alpha - 1} \log \left( (\rho^*)^{\frac{1}{\alpha}} \frac{\rho^* + 2\phi}{\phi\rho^* + 1 + \phi^2} \right). \tag{J.2}$$

Since $\rho^* > \rho_L$ under strong increasing returns, the expression under the log is greater than 1 and thus the RHS of (J.2) is negative. Furthermore, as $\rho^*$ decreases with $\gamma$, the RHS of (J.2) increases with $\gamma$. In addition, the first of the RHS goes to 0 when $\gamma$ goes to infinity. Consequently, $\ell^*$ increases and converges to 1. Q.E.D.

### *K. Proof of Lemma 4*

The assumption $s^*_{ij} > 0$ implies $L^*_j > 0$, hence $W^*_j > 0$. Combining this with (2) and (34) implies that any individual $\nu \in [0,1]$ whose type $\mathbf{z}(\nu)$ satisfies

$$z_{ij}(\nu) = z_{kl}(\nu)\frac{W^*_l}{W^*_j}\frac{\tau_{kl}}{\tau_{ij}}\left(\frac{R^*_i}{R^*_k}\right)^{\beta} \geq z_{od}(\nu)\frac{W^*_d}{W^*_j}\frac{\tau_{od}}{\tau_{ij}}\left(\frac{R^*_o}{R^*_i}\right)^{\beta} \tag{K.1}$$

is indifferent between $ij$ and $kl$.

Two cases may arise. First, if $s^*_{kl} > 0$, then $L^*_l > 0$ and $W^*_l > 0$. (H.1) thus implies that any individual $\nu$ satisfying

$$z_{kl}(\nu) > 0, \quad z_{ij}(\nu) = z_{kl}(\nu)\frac{W^*_l}{W^*_j}\frac{\tau_{kl}}{\tau_{ij}}\left(\frac{R^*_i}{R^*_k}\right)^{\beta}, \quad z_{od}(\nu) = 0$$

is indifferent between $ij$ and $kl$.

Second, if $s^*_{kl} = 0$, then $L^*_l = 0$ and $W^*_l = 0$. Therefore, (K.1) implies that any individual such that $z_{kl}(\nu) > 0$ and $z_{ij}(\nu) = 0$ for any $ij \neq kl$ is indifferent between $ij$ and $kl$. Q.E.D.

### *L. Proof of Proposition 9*

**Step 1.** We first show the existence of a unique conditional equilibrium price for a symmetric commuting pattern $\mathbf{s}$ such that either $\mathbf{L}(\mathbf{s}) = (0,1,0)$ or $\mathbf{L}(\mathbf{s}) = (1/2,0,1/2)$, and $M_i(\mathbf{s}) > 0$ for $i = 0, \pm 1$.

We focus on the case of a fully agglomerated labor supply pattern, i.e., such that $L_0(\mathbf{s}) = 1$ and $L_{-1}(\mathbf{s}) = L_1(\mathbf{s}) = 0$ (the proof for the fully dispersed labor supply pattern given by $L_0 = 0$ and $L_{-1}(\mathbf{s}) = L_1(\mathbf{s}) = 1/2$ goes along the same lines). Plugging $L_0 = 1$ and $L_{-1} = L_1 = 0$ into the FOCs (13) and (14) at $i = 0$, we obtain

$$W_0 = \alpha N_0^{1-\alpha} \qquad \text{and} \qquad R_0 = (1-\alpha) N_0^{-\alpha}, \tag{L.1}$$

hence,

$$\frac{W_0}{R_0} = \frac{\alpha}{1-\alpha} N_0. \tag{L.2}$$

Observe that $L_1(\mathbf{s}) = L_{-1}(\mathbf{s}) = 0$ implies $s_{i1} = s_{i,-1} = 0$ for all $i \in \{-1,0,1\}$. Combining this with the land market clearing condition and the market residential demand at $i = 0$, we get:

$$H_0 + N_0 = 1 \qquad \text{and} \qquad H_0 = s_{00} \frac{W_0}{R_0},$$

so that

$$N_0 = 1 - H_0 = 1 - s_{00} \frac{W_0}{R_0}. \tag{L.3}$$

Plugging (L.3) into (L.2), we get a linear equation in $W_0/R_0$:

$$\frac{W_0}{R_0} = \frac{\alpha}{1-\alpha} \left( 1 - s_{00} \frac{W_0}{R_0} \right) \quad \Longrightarrow \quad \frac{\overline{W}_0(\mathbf{s})}{\overline{R}_0(\mathbf{s})} = \frac{\alpha}{1-\alpha+\alpha s_{00}}. \tag{L.4}$$

From (L.3)-(L.4), we get:

$$\overline{N}_0(\mathbf{s}) = \frac{1-\alpha}{1-\alpha+\alpha s_{00}}.$$

Plugging $N_0 = \overline{N}_0(\mathbf{s})$ into the equilibrium condition (L.1) pins down uniquely the conditional equilibrium wage $\overline{W}_0(\mathbf{s})$ and the conditional equilibrium land rent $\overline{R}_0(\mathbf{s})$. As for $\overline{W}_j(\mathbf{s})$ and $\overline{R}_i(\mathbf{s})$ for $i,j = \pm 1$, zero labor supply implies $\overline{W}_j(\mathbf{s}) = 0$ and $\overline{N}_j(\mathbf{s}) = 0$ for $j = \pm 1$. Hence, the land market clearing at the periphery becomes

$$H_i = 1 = s_{i0} \frac{W_0}{R_i} \qquad \text{for} \qquad i = \pm 1,$$

which implies $\overline{R}_i(\mathbf{s}) = s_{i0} \overline{W}_0(\mathbf{s})$ for $i = \pm 1$.

**Step 2.** We now show that corner equilibria are unstable. Assume that $L_0^* = 1$ (the proof for $L_{-1}^* = L_1^* = 1/2$ goes along the same lines). Consider an individual $\nu$ such that, for all $i \in \{-1,0,1\}$, $\nu$'s match values satisfy $z_{ij}(\nu) = 0$ for $j = 0, \pm 1$. Clearly, $\nu$ is indifferent between working at the center and working at the periphery (in both cases, she enjoys zero utility). Consider a positive-measure set of individuals whose tastes are close to those of $\nu$ and whose utility-maximizing choice is $ij = 00$. Relocating them (together with $\nu$) from $ij = 00$ to $kl = 01$, we have $V_{01}(\nu,\mathbf{s}) > 0$ because $\overline{W}_1(\mathbf{s}) > 0$. Using (13), there exists a positive-measure subset of individuals who are strictly better-off working at location 1. As a result, the corner equilibrium $L_0^* = 1$ is an unstable equilibrium. Q.E.D.

***M. Proof of Proposition 10***

The proof involves 4 steps.

**Step 1.** We first show the existence of a unique conditional equilibrium price for a symmetric commuting $\mathbf{s}$ such that $s_{ij} > 0$ for all $i,j$ when $\alpha > 1/2$.

Since $L_i > 0$ for $i = 0, \pm 1$, the first-order conditions for the production sector yields the equilibrium conditions:

$$W_j = \alpha A_j \left(\frac{N_j}{L_j}\right)^{1-\alpha}, \tag{M.1}$$

$$R_j = (1-\alpha) A_j \left(\frac{L_j}{N_j}\right)^{\alpha}. \tag{M.2}$$

Furthermore, we also know that housing market clearing at location $i$ yields:

$$H_i = \frac{\beta}{R_i} \sum_{j=1}^{n} s_{ij} W_j. \tag{M.3}$$

Plugging (M.1) and (M.2) into (M.3), and using the land market balance condition $N_i + H_i = 1$, we get:

$$H_i = 1 - N_i = \frac{\alpha\beta}{(1-\alpha)A_i} \left(\frac{N_i}{L_i}\right)^{\alpha} \sum_{j=1}^{n} s_{ij} A_j \left(\frac{N_j}{L_j}\right)^{1-\alpha},$$

$$(1-\alpha) A_i (1 - N_i) \left(\frac{N_i}{L_i}\right)^{-\alpha} = \alpha\beta \sum_{j=1}^{n} s_{ij} A_j \left(\frac{N_j}{L_j}\right)^{1-\alpha},$$

$$(1-\alpha) A_i \left(\frac{N_i}{L_i}\right)^{-\alpha} = (1-\alpha) A_i L_i \left(\frac{N_i}{L_i}\right)^{1-\alpha} + \alpha\beta \sum_{j=1}^{n} s_{ij} A_j \left(\frac{N_j}{L_j}\right)^{1-\alpha}.$$

Since **s** is symmetric, this system of equations becomes:

$$(1-\alpha)A_0\left(\frac{N_0}{L_0}\right)^{-\alpha} = \left[(1-\alpha)L_0+\alpha\beta s_{00}\right]A_0\left(\frac{N_0}{L_0}\right)^{1-\alpha} + 2\alpha\beta s_{01}A_1\left(\frac{N_1}{L_1}\right)^{1-\alpha}$$

$$(1-\alpha)A_1\left(\frac{N_1}{L_1}\right)^{-\alpha} = \alpha\beta s_{10}A_0\left(\frac{N_0}{L_0}\right)^{1-\alpha} + \left[(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})\right]A_1\left(\frac{N_1}{L_1}\right)^{1-\alpha}$$

Dividing one equation by the other and using $A_i = L_i^{\gamma}$ for $i = 0, \pm 1$, we get:

$$n^{-\alpha}\ell^{\gamma+\alpha} = \frac{\left[(1-\alpha)L_0+\alpha\beta s_{00}\right]\ell^{\gamma}\left(\frac{n}{\ell}\right)^{1-\alpha}+2\alpha\beta s_{01}}{\alpha\beta s_{10}\ell^{\gamma}\left(\frac{n}{\ell}\right)^{1-\alpha}+(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})} \tag{M.4}$$

Since (M.1) and (M.2) imply

$$n^{-\alpha}\ell^{\gamma+\alpha} = r, \qquad \ell^{\gamma}\left(\frac{n}{\ell}\right)^{1-\alpha} = w, \tag{M.5}$$

we have

$$w^{\alpha}r^{1-\alpha} = \ell^{\gamma} = \left(\frac{L_0}{L_1}\right)^{\gamma}. \tag{M.6}$$

Likewise, combining (M.4) and (M.5), we get:

$$r = \frac{\left[(1-\alpha)L_0+\alpha\beta s_{00}\right]w+2\alpha\beta s_{01}}{\alpha\beta s_{10}w+(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})}. \tag{M.7}$$

A sufficient condition for the system (M.6) – (M.7) to have a unique solution $(\overline{w}(\mathbf{s}),\overline{r}(\mathbf{s}))$ is that the graph of the relationship (M.7) between $w$ and $r$ intersects the downward-sloping curve given by (M.6) from below. The RHS of (M.7) is the ratio of two positive linear increasing functions of $w$. Since the elasticity of a linear increasing function with a positive intercept never exceeds 1, the elasticity of the RHS of (M.7) w.r.t. $w$ is always larger than $-1$. Restating (M.6) as

$$r = \ell^{\frac{\gamma}{1-\alpha}}w^{-\frac{\alpha}{1-\alpha}}$$

shows that the elasticity of the RHS of this expression w.r.t. $w$ equals $-\alpha/(1-\alpha)$, which is smaller than $-1$ when $\alpha > 1/2$.

**Step 2.** Denote by $\left(\overline{\mathbf{W}}(\mathbf{s}),\overline{\mathbf{R}}(\mathbf{s})\right)$ the equilibrium price vector conditional to an arbitrary commuting pattern **s** that belongs to a neighborhood of an interior equilibrium commuting pattern $\mathbf{s}^*$, and let $\overline{w}(\mathbf{s})$ and $\overline{r}(\mathbf{s})$ be the corresponding wage ratio and the land-price ratio:

$$\overline{w}(\mathbf{s}) \equiv \frac{\overline{W}_0(\mathbf{s})}{\overline{W}_1(\mathbf{s})} \quad \text{and} \quad \overline{r}(\mathbf{s}) \equiv \frac{\overline{R}_0(\mathbf{s})}{\overline{R}_1(\mathbf{s})}.$$

Consider the following two types of relocations: $0j \to 1j$ (changing place of residence but not the workplace) and $i0 \to i1$ (changing the workplace but not the place of residence). Observe that, in equilibrium, for each individual $\nu$, we have:

$$\frac{V_{0j}^*(\nu)}{V_{1j}^*(\nu)} = \frac{z_{0j}(\nu)}{z_{1j}(\nu)} \left(\overline{r}(\mathbf{s}^*)\right)^{-\beta}, \tag{M.8}$$

$$\frac{V_{i0}^*(\nu)}{V_{i1}^*(\nu)} = \frac{z_{i0}(\nu)}{z_{i1}(\nu)} \overline{w}(\mathbf{s}^*). \tag{M.9}$$

If the individual $\nu$ is indifferent between $0j$ and $1j$ for some $j = \{-1,0,1\}$, switching from $0j$ to $1j$ makes this individual strictly worse off if and only if $\overline{r}(\mathbf{s}^*)$ decreases when a small subset of residents (almost indifferent between $0j$ and $1j$) of measure $\Delta$ is moved from 0 to 1, i.e.,

$$\frac{\partial \overline{r}(\mathbf{s}^*)}{\partial s_{1j}} - \frac{\partial \overline{r}(\mathbf{s}^*)}{\partial s_{0j}} < 0 \tag{M.10}$$

because (M.8) and (M.10) imply that $V_{0j}^*(\nu)/V_{1j}^*(\nu)$ increases above 1.

Likewise, using (M.9) if $\nu$ is an individual indifferent between $i0$ and $i1$ for some $i = \{-1,0,1\}$, switching from $i0$ to $i1$ makes $\nu$ strictly worse off if and only if $\overline{w}(\mathbf{s}^*)$ increases when a small subset of workers (almost indifferent between $i0$ and $i1$) of measure $\Delta$ is moved from 0 to 1, i.e.,

$$\frac{\partial \overline{w}(\mathbf{s}^*)}{\partial s_{i1}} - \frac{\partial \overline{w}(\mathbf{s}^*)}{\partial s_{i0}} > 0. \tag{M.11}$$

**Step 3.** We now show that the land-price ratio $\overline{r}(\mathbf{s}^*)$ always satisfies the equilibrium condition (M.10). Under a relocation of residents from $0j$ to $1j$ (or, equivalently, from $1j$ to $0j$) for $j = 0,1$, the numerator in the RHS of (M.7) decreases pointwise, while the denominator increases pointwise. Therefore, the curve (M.7) shifts downwards in the $(w,r)$-plane, while the curve (M.6) remains unchanged. Since (M.7) intersects (M.6) from below, this implies a reduction in $\overline{r}(\mathbf{s})$. Hence, (M.10) holds.

**Step 4.** It remains to check when (M.11) holds. To this end, we study when the relocation of a $\Delta-$measure subset of workers from $i0$ to $i1$ for $i = 0, \pm 1$ leads to an increase in the relative wage $\overline{w}(\mathbf{s})$. As a result, two cases must be distinguished: (i) a relocation of workers from 00 to 01 and (ii) a relocation of workers from 10 to 11.

Taking the log-differential of (M.6) yields:

$$\alpha d \log w + (1-\alpha) d \log r = \gamma (d \log L_0 - d \log L_1). \tag{M.12}$$

*Case 1.* Assume that

$$ds_{00} = -\Delta, \qquad ds_{01} = ds_{0,-1} = \Delta/2,$$

$$ds_{ij} = 0 \text{ otherwise.}$$

In this case, (M.12) becomes:

$$\alpha\, d \log w + (1-\alpha)\, d \log r = \gamma \left( \frac{ds_{00}}{L_0} - \frac{ds_{01}}{L_1} \right) = -\gamma \Delta \left( \frac{1}{2L_1} + \frac{1}{L_0} \right)$$

Taking the log-differential of (M.7) yields:

$$d \log r = \frac{d\left[\left((1-\alpha)L_0 + \alpha\beta s_{00}\right) w + 2\alpha\beta s_{01}\right]}{\left[(1-\alpha)L_0 + \alpha\beta s_{00}\right] w + 2\alpha\beta s_{01}} - \frac{d\left[\alpha\beta s_{10} w + (1-\alpha)L_1 + \alpha\beta(s_{11} + s_{1,-1})\right]}{\alpha\beta s_{10} w + (1-\alpha)L_1 + \alpha\beta(s_{11} + s_{1,-1})}. \tag{M.13}$$

Since

$$d\left[\left((1-\alpha)L_0 + \alpha\beta s_{00}\right) w + 2\alpha\beta s_{01}\right] = -\Delta\left(1-\alpha+\alpha\beta\right) w + \alpha\beta\Delta + \left((1-\alpha)L_0 + \alpha\beta s_{00}\right) w d \log w,$$

while

$$d\left[\alpha\beta s_{10} w + (1-\alpha)L_1 + \alpha\beta(s_{11} + s_{1,-1})\right] = (1-\alpha)\frac{\Delta}{2} + \alpha\beta s_{10} w d \log w,$$

(M.13) becomes

$$d \log r = \left[ \frac{-\left(1-\alpha+\alpha\beta\right) w + \alpha\beta}{\left((1-\alpha)L_0 + \alpha\beta s_{00}\right) w + 2\alpha\beta s_{01}} - \frac{1}{2} \frac{1-\alpha}{\alpha\beta s_{10} w + (1-\alpha)L_1 + \alpha\beta(s_{11} + s_{1,-1})} \right] \Delta$$

$$+\left[\frac{((1-\alpha)L_0+\alpha\beta s_{00})\,w}{((1-\alpha)L_0+\alpha\beta s_{00})\,w+2\alpha\beta s_{01}}-\frac{\alpha\beta s_{10}w}{\alpha\beta s_{10}w+(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})}\right]d\log w$$

Plugging this expression into (M.12), we get:

$$d\log w=\frac{-\gamma\left(\frac{1}{2L_1}+\frac{1}{L_0}\right)+(1-\alpha)\left[\frac{(1-\alpha+\alpha\beta)w-\alpha\beta}{((1-\alpha)L_0+\alpha\beta s_{00})w+2\alpha\beta s_{01}}+\frac{1}{2}\frac{1-\alpha}{\alpha\beta s_{10}w+(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})}\right]}{\alpha+(1-\alpha)\left[\frac{((1-\alpha)L_0+\alpha\beta s_{00})w}{((1-\alpha)L_0+\alpha\beta s_{00})w+2\alpha\beta s_{01}}-\frac{\alpha\beta s_{10}w}{\alpha\beta s_{10}w+(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})}\right]}\Delta.$$

When $\alpha>1/2$, the denominator in $d\log w$ is always positive because each bracketed term of the denominator is smaller than 1. As a result, the stability condition $d\log w>0$ holds if the numerator is positive:

$$\frac{(1-\alpha+\alpha\beta)\,w-\alpha\beta}{((1-\alpha)L_0+\alpha\beta s_{00})\,w+2\alpha\beta s_{01}}+\frac{1}{2}\frac{1-\alpha}{\alpha\beta s_{10}w+(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})}>\frac{\gamma}{1-\alpha}\left(\frac{1}{L_0}+\frac{1}{2L_1}\right). \tag{M.14}$$

*Case 2*. We now assume that

$$ds_{11}=-ds_{10}=\Delta/2,\qquad ds_{-10}=-ds_{-1,-1}=-\Delta/2,$$

$$ds_{ij}=0\text{ otherwise.}$$

Hence, (M.12) becomes:

$$\alpha d\log w+(1-\alpha)\,d\log r=\gamma\left[\frac{ds_{10}+ds_{-10}}{L_0}-\frac{ds_{11}}{L_1}\right]=-\gamma\Delta\left(\frac{1}{2L_1}+\frac{1}{L_0}\right)$$

Since

$$d\left[((1-\alpha)L_0+\alpha\beta s_{00})\,w+2\alpha\beta s_{01}\right]=-\Delta\,(1-\alpha)\,w+((1-\alpha)L_0+\alpha\beta s_{00})\,wd\log w$$

and

$$d\left[\alpha\beta s_{10}w+(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})\right]=\alpha\beta\frac{\Delta}{2}w+(1-\alpha)\frac{\Delta}{2}+\alpha\beta s_{10}wd\log w,$$

(M.13) becomes

$$d\log r = \left[\frac{-(1-\alpha)\,w}{((1-\alpha)L_0+\alpha\beta s_{00})\,w+2\alpha\beta s_{01}} - \frac{1}{2}\frac{1-\alpha+\alpha\beta}{\alpha\beta s_{10}w+(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})}\right]\Delta$$

$$+\left[\frac{((1-\alpha)L_0+\alpha\beta s_{00})\,w}{((1-\alpha)L_0+\alpha\beta s_{00})\,w+2\alpha\beta s_{01}} - \frac{\alpha\beta s_{10}w}{\alpha\beta s_{10}w+(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})}\right]d\log w.$$

Plugging this expression for $d\log r$ into (M.12), we get:

$$d\log w \;=\frac{-\gamma\left(\frac{1}{2L_1}+\frac{1}{L_0}\right)+(1-\alpha)\left[\frac{(1-\alpha)w}{((1-\alpha)L_0+\alpha\beta s_{00})w+2\alpha\beta s_{01}}+\frac{1}{2}\frac{1-\alpha+\alpha\beta}{\alpha\beta s_{10}w+(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})}\right]}{\alpha+(1-\alpha)\left[\frac{((1-\alpha)L_0+\alpha\beta s_{00})w}{((1-\alpha)L_0+\alpha\beta s_{00})w+2\alpha\beta s_{01}}-\frac{\alpha\beta s_{10}w}{\alpha\beta s_{10}w+(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})}\right]}\Delta.$$

If $\alpha > 1/2$, the denominator in $d\log w$ is always positive. Hence, the stability condition $d\log w > 0$ becomes:

$$\frac{(1-\alpha)\,w}{((1-\alpha)L_0+\alpha\beta s_{00})\,w+2\alpha\beta s_{01}}+\frac{1}{2}\frac{1-\alpha+\alpha\beta}{\alpha\beta s_{10}w+(1-\alpha)L_1+\alpha\beta(s_{11}+s_{1,-1})} > \frac{\gamma}{1-\alpha}\left(\frac{1}{L_0}+\frac{1}{2L_1}\right). \tag{M.15}$$

When $\alpha > 1/2$, the inequalities (M.14) and (M.15) are necessary and sufficient for an interior equilibrium to be stable.

We now rewrite these two conditions in terms of the variable $\rho$ only. Using Proposition 1 and the equilibrium relationship $\omega^{\frac{1+\varepsilon}{\varepsilon}} = f(\rho)$, as well as $\rho = r^{-\beta\varepsilon}$, $\omega = w^{\varepsilon}$, and $a = \alpha\beta/(1-\alpha)$, (M.14) and (M.15) become

$$\frac{f(\rho)+\frac{1}{2}(1+a)\rho^{-\frac{1}{\beta\varepsilon}}\left[f(\rho)\right]^{\frac{\varepsilon}{1+\varepsilon}}}{((1+a)\rho+2\phi)\,f(\rho)+2a\phi\rho} > \frac{\gamma}{1-\alpha}\left(\frac{\left[f(\rho)\right]^{\frac{\varepsilon}{1+\varepsilon}}}{2(\phi\rho+1+\phi^2)}+\frac{1}{\rho+2\phi}\right), \tag{M.16}$$

$$\frac{(1+a)\,f(\rho)+\left(\frac{1}{2}\rho^{-\frac{1}{\beta\varepsilon}}-a\right)\left[f(\rho)\right]^{\frac{\varepsilon}{1+\varepsilon}}}{((1+a)\rho+2\phi)\,f(\rho)+2a\phi\rho} > \frac{\gamma}{1-\alpha}\left(\frac{\left[f(\rho)\right]^{\frac{\varepsilon}{1+\varepsilon}}}{2(\phi\rho+1+\phi^2)}+\frac{1}{\rho+2\phi}\right), \tag{M.17}$$

Solving the equilibrium condition $f(\rho) = g(\rho;\gamma)$ w.r.t. $\gamma$ yields

$$\gamma = \frac{\alpha}{1+\varepsilon}\frac{\log\left(\rho^{-b}f(\rho)\right)}{\log\left(\frac{\rho+2\phi}{\phi\rho+1+\phi^2}\left[f(\rho)\right]^{\frac{\varepsilon}{1+\varepsilon}}\right)}.$$

Plugging this expression into (M.16) – (M.17), we get:

$$\Phi(\rho) \equiv \frac{2(1-\alpha)(1+\varepsilon)(\phi\rho+1+\phi^2)}{((1+a)\rho+2\phi)\, f(\rho)+2a\phi\rho} \times \frac{f(\rho)+\left(\frac{1}{2}\rho^{-\frac{1}{\beta\varepsilon}}+\frac{a}{2}\rho^{-\frac{1}{\beta\varepsilon}}\right)[f(\rho)]^{\frac{\varepsilon}{1+\varepsilon}}}{[f(\rho)]^{\frac{\varepsilon}{1+\varepsilon}}(\rho+2\phi)+2(\phi\rho+1+\phi^2)} \times \frac{\log\left(\frac{\rho+2\phi}{\phi\rho+1+\phi^2}[f(\rho)]^{\frac{\varepsilon}{1+\varepsilon}}\right)}{\alpha\log\left(\rho^{-b}f(\rho)\right)} > 1,$$

$$\Psi(\rho) \equiv \frac{2(1-\alpha)(1+\varepsilon)(\phi\rho+1+\phi^2)}{((1+a)\rho+2\phi)\, f(\rho)+2a\phi\rho} \times \frac{(1+a)\, f(\rho)+\left(\frac{1}{2}\rho^{-\frac{1}{\beta\varepsilon}}-a\right)[f(\rho)]^{\frac{\varepsilon}{1+\varepsilon}}}{[f(\rho)]^{\frac{\varepsilon}{1+\varepsilon}}(\rho+2\phi)+2(\phi\rho+1+\phi^2)} \times \frac{\log\left(\frac{\rho+2\phi}{\phi\rho+1+\phi^2}[f(\rho)]^{\frac{\varepsilon}{1+\varepsilon}}\right)}{\alpha\log\left(\rho^{-b}f(\rho)\right)} > 1.$$

Last, we set:

$$\mathbb{F}(\rho) \equiv \min\left\{\Phi(\rho),\Psi(\rho)\right\},$$

which is independent of $\rho$. Verifying $\mathbb{F}(\rho) > 1$ can be done numerically for any vector of parameters by plotting $\mathbb{F}(\rho)$ as a function of the variable $\rho$. Q.E.D.