Comparative-Static Heuristics in Spatial Models

Mark Bamba*

November 12, 2025
Latest draft: link

Abstract

We consider first-order comparative statics in constant-elasticity spatial models. In this class, exact counterfactual and first-order comparative-static analysis have the same structure and common data requirements: model elasticities and the matrices of inter-location flows for each interaction, as emphasized in the international trade and spatial literatures. These comparative statics can be decomposed into a direct effect and a general-equilibrium (GE) remainder. We show that the relative magnitudes of these two components can be characterized using the spectral radius and matrix norms of a matrix constructed from the elasticities.

We argue that, in this family of models, both first-order and exact comparative statics can be analyzed using these direct effects, which we term "heuristics": closed-form expressions in terms of observables and parameters that (i) require fewer inputs, (ii) are more compact, and (iii) illuminate the mechanisms that matter most in a given specification. For first-order comparative statics, we provide theoretical results that quantify the size of the heuristic relative to the GE remainder—using similar inputs needed to compute the heuristic itself. These bounds allow researchers to make sharp predictions about the direction of comparative statics without solving the full GE system, provided GE forces implied by the elasticity matrix are not too strong.

We apply these tools to gravity-based urban commuting models to study how transport systems shape residence and employment. The theory yields new insights into the effects of radial versus core-focused (e.g., center-city subway) infrastructure on the spatial distribution of employment and population. Implementing the results for Tokyo's rail network, we obtain sharp predictions for the sign and ranking of population and employment changes from transport improvements—without computing the GE remainder—while emphasizing that computing the GE component remains important for accurate level predictions and full welfare evaluation.

*We thank Eduardo Morales, Yuhei Miyauchi, David Weinstein, Richard Rogerson, Gene Grossman, John Sturm, John Sturm, Michael Jenuwine, Yuyang Jiang among others. Participants from the Princeton Student Trade Workshop are thanked for their constructive comments.

Introduction

We consider comparative statics for a large class of spatial models, which we refer to as constant elasticity spatial models. For this class of models, we analyze comparative statics that describe the change in endogenous variables due to a change in geographically-related parameters (e.g. fundamental productivity of a location, transportation costs between two locations, etc.). In this class, exact counterfactual and first-order comparative-static analysis have the same structure and common data requirements: model elasticities and the matrices of inter-location flows for each interaction, as emphasized in the international trade and spatial literatures. Specifically, to determine the change in endogenous variables for a change in geographically-related parameters, all the researcher needs are the elasticities and the matrix of "flow shares", which are the fraction of flows entering (or exiting) one location that is coming from (or going to) another location.

In the case of first-order analysis, the comparative static change in endogenous variables can be decomposed into a partial equilibrium effect, which may be expressed explicitly and succinctly using observables, and the remaining general equilibrium effect. We show that the partial-equilibrium effect is the leading (and therefore largest) term of a Neumann series representation of the general equilibrium effect whenever spatial spillovers are not too large, and we can therefore use the decaying rate of the series (as measured by the spectral radius or matrix norms of what we term the "spillover matrix") to understand the relative sizes of the partial equilibrium and remaining general equilibrium effects.

Depending on the researcher's needs, one can adjust the definition of the partial equilibrium effect by redefining what variables are held constant. This change in definition alters the partial-equilibrium expression and the quantitative size of the partial equilibrium effect relative to the general equilibrium remainder. The relative sizes can once again be analyzed using the new spillover matrix. The researcher may thus wish to consider various definitions of the "partial equilibrium effect," depending on the empirical setting and the need for accuracy and theoretical insight.

In exact comparative static analysis, one can also construct an *exact* partial equilibrium expression that can be derived from the system of equations and that is analagous to that of the first-order partial equilibrium analysis. While we do not characterize the theoretical properties of the exact partial-equilibrium expression relative to the general equilibrium expression, we note that the exact partial-equilibrium expression may have more relevance and predictive power when studying large comparative statics changes.

We then apply these tools to a class of gravity-based urban commuting models to understand the mechanisms underlying the impact of transportation systems on where people live and work. The theory provides new insights into how radial versus centralized commuting infrastructure affects residential location choices. We focus on residential population, as this was the surprising result in our empirical application, but analogous arguments can be made for employment.

Focusing on how symmetric transportation improvements (e.g. when commuting costs fall by the same amount in both directions) between two locations affect population in those locations, we show that the fraction of each location's residents who commute to the other location are sufficient statistics for the partial equilibrium change in relative population between the two locations, where relative employment are held constant in all locations.[1]

Specifically, the difference in these two statistics between two locations is proportional to the partial equilibrium effect of a marginal symmetric transport improvement between the two locations on their relative populations. In other words, if a higher proportion of residents from one location commute to the other location than vice versa, then the location with the higher proportion gains population relative to the other location according to the partial-equilibrium effect.

Our earlier results imply that this difference represents the leading term of a decaying Neumann series that expresses the full GE effect, and we analyze the quantitative importance and usefulness of this sufficient statistic as a proxy for the general equilibrium comparative static in two ways.

First, we use our theory to compute bounds on the remaining GE effect based on elasticities estimated in the literature, and provide conditions that guarantee the direction of relative population changes between two locations. When we consider all symmetric improvements in commuting infrastructure between pairs of municipalities in Tokyo (all in isolation), we find that in 44% of improvements our sufficient conditions are satisfied and we can guarantee a population increase one way or the other. Using the same exercise of isolated pair-wise symmetric improvements, we then show empirically that this sufficient statistic accounts for more than 97% of the variation in the first-order general equilibrium effect in the case of Tokyo and exhibits little statistical bias (less than 3% of the GE effect) when the GE effect is regressed on the sufficient statistic.

The reason for the strong predictions and high predictability lies in the fact that our spatial spillovers are relatively mild, leading to a spectral radius of 0.42, which, a back-of-the-envelope calculation shows, translates to the partial equilibrium effect accounting for at least 56% of the full general equilibrium (first-order) effect for the relative population for

[1] "Relative" employment refers to employment relative to a hypothetical location that is fully isolated from the rest of the economy–where there is no commuting into or out of this location. Thus even as relative employment is held constant, the *level* of employment may change depending on population mobility constraints of the model.

any individual location.

What are the implications for understanding the impact of commuting infrastructure improvements on the location of residential population? The implications of these results are twofold. First, radial improvements in commuting infrastructure (i.e., those radiating from a central location to a peripheral location) tend to decentralize population away from the center whenever commuting from the suburb to downtown is more common than "reverse" commuting (e.g., from the center to a suburb), as is typical in cities.

In the Tokyo context, the average municipality in the periphery has a 20 times higher share of commuters to the center than the other way around. Thus, radial improvements in Tokyo will tend to shift population to the periphery if we consider transport costs between locations in isolation.[2] A caveat to our analysis is that transportation infrastructure forms a network. Therefore, a change in one link of the network changes commute costs across many pairs in non-uniform ways. We refer the reader to Allen and Arkolakis (2022), which provides an elegant methodology to study this problem.

Second, subways and 'central'-focused transportation improvements may improve transportation for all commuters but disproportionately benefit residents living in the center, leading to centralization of population in the center. Consider, for example, a symmetric 4-location model on a line with a "periphery" location at each end, each of which is in turn connected to two central "core" locations, which are themselves connected. Locations are indexed from 1 to 4, with locations 2 and 3 representing the core and locations 1 and 4 the periphery. The connections form a linear chain: 1–2–3–4.

Suppose, for simplicity, that there is no commuting across the city so that locations 1 and 4 cannot reach each other. Now suppose that commuting infrastructure improves along the link between the two cores, such that commuting costs fall between the two cores and between the peripheries and the cores. Under common parameterizations of commuting costs, the percentage drop in commuting costs between the core and periphery is at least as large as the drop between the two cores. Denoting d_{ij} as a measure of commuting costs for commutes from residents in i to workplace location j, models typically assume that $|d \ln d_{23}| \geq |d \ln d_{13}|$.[3] In such a case, the transportation improvement will tend to shift population from the periphery

[2] Our bounds indicate that symmetric improvements between peripheral locations and the center are guaranteed to lead to an increase in peripheral municipality population for 64% of the pairs of locations considered, while the central population increases about 1% of the time.

[3] For example, if commuting costs are the product across links: $d_{13} = d_{12} d_{23}$, then $d \ln d_{13} = d \ln d_{23}$. If they are a power function of total commute time, $d(t) = t^{\xi}$, then $|d \ln d_{13}| = \left| \frac{\xi t_{23}}{t_{12} + t_{23}} d \ln t_{23} \right| < |\xi d \ln t_{23}| = |d \ln d_{23}|$. More generally, if commuting costs are a function of total commute time $d = f(t)$ where t is total commute time, then this condition is violated (i.e., $|d \ln d_{13}| > |d \ln d_{23}|$) if the elasticity of commuting costs grows faster than linearly in travel time: $\frac{d \ln}{d \ln t} \frac{d \ln f}{d \ln t} > 1$. One example is the function $f(t) = e^{t^2}$, a functional form that is rarely used and not supported by Tokyo's commuter-flow data.

to the centers if the proportion of core populations that commute to each other is greater than the proportion of peripheral populations that commute to the further-away core location (commuting from 1 to 3 or from 4 to 2).

In the Tokyo context, relative to the average peripheral municipality (which we define as the region more than 40 km from the Imperial Palace), the average central municipality (the region within 20 km of the Imperial Palace) is 26 times more likely to commute within central Tokyo, and the average suburban municipality (defined as the ring between 20–40 km from the Imperial Palace) is 5 times more likely to commute to central Tokyo. Our theory indicates that these large disparities will tend to shift population to the center when there are central-focused transportation improvements.

Last, we apply these tools to study the impact of Tokyo's train system on residential population across the city.

The full exact-hat quantitative results show that Tokyo's train system leads to centralization in both employment and population. In other words, the model implies that the system causes a shift in population from the peripheries and suburbs to the city center. These results contrast with what researchers have often found for other commuting infrastructures in other cities and raise the question of the mechanisms of the model, something we explore as our final contribution.

The first-order partial equilibrium effect on each residential location is proportional to a weighted average of changes in our measure of accessibility (i.e., the inverse of transformed commute costs) between the location and all other locations, where the weights are the commuting shares to each location from the residential location. The exact partial equilibrium can be expressed similarly.

Regression of the first-order general equilibrium effect on the first-order partial-equilibrium effect shows that the partial equlibrium effect explain 82% of the variation in the general equlibrium first-order population change due to Tokyo's train system. The slope coefficient, however, is 0.67, indicating that the partial-equilibrium expression is biased downwards due to the congetion forces that are embodied in the two spillover elasticities, which are both negative. Regression of the exact general equilibrium effect on the exact partial equilibrium effect has an R^2 of 0.91 and an OLS coefficient of 0.6. These results indicate that the partial equilibrium effect measures have large predictive power (under our choice of parameters), but that their numerical quantity alone should not be directly used at face value for comparative-static analysis.

Given our theoretical results and strong predictability between the partial-equilibrium and general equilibrium effect, we suspect that a major factor for the centralizing effect of Tokyo's train system is that central Tokyo, has significantly higher commuting access (i.e.

lower commuting costs) than the suburbs and peripheries, due to the dense train network in the center. Indeed, the average central municipality gains 1.9 times more in residential access in partial equilibrium, relative to the periphery. The results from the present paper argues that this disparity tends to a large population shift from the periphery to the center.

In our final quantification, we show quantitatively that population centralization in Tokyo is (as we predicted from our theory) largely due to the dense subway network that is in the urban core. In counterfactuals we remove segments of the subway lines one by one in the reverse order in which they were built. Consistent with theory, we find that each subway segment was responsible for the increase in the population in the core. In contrast, when we do the same exercisee for non subway lines, which are more often commuter lines that radiate the suburbs, we find that these lines tended to shift the population to the peripheries. Thus, our counterfactuals corroborate our theoretical predictions that central-focused transportation improvements have a centralizing effect on the city while our theoretical results provides an explanation for why the dense subway network in the urban core causes population centralization.

The centralizing impact of transportation infrastructure on population has few empirical precedents. Previous studies have often found centralizing impacts of new transportation on employment but rarely on population. Tokyo therefore stands as an outlier. At the same time, Tokyo is one of the most densely populated cities in the world (see for example, table 1) with a population of approximately 40 million, roughly twice the population of the New York-Newark metropolitan area, contained within an area that is about 30% smaller. Our insights, both theoretical and using Tokyo data, help us understand not only what facilitates Tokyo's dense population but also other cities' which happen to have large train networks and high central population densities, such as New York City, Paris, London, Hong Kong, Shanghai, and Seoul among others.

1 Literature

Our theoretical work on spatial models is closely related to three papers. Allen et al. (2024) characterize the uniqueness properties of the class of models studied in the present paper and focus on the matrix of elasticities that the present paper also emphasizes. They also provide theoretical methods (change of variables, absorbing endogenous scalars) that are used in the present paper.

Allen et al. (2020) provide comparative static expressions and Neumann series expansions for a wide class of constant-elasticity trade models (e.g., Arkolakis et al. (2012)), which the authors refer to as "universal gravity" models. These models (similar to the commuting model used in our study) are characterized by two primary interactions—one representing the recipient and one representing the sender. In the trade models, the recipient is the importer and the sender is the exporter, while in commuting models, the recipient is the workplace and the sender is the residential location.

In Allen and Arkolakis (2025), as well as reviewing the variety of methods in economic geography models, the authors use comparative static methods under the "universal gravity" framework and apply them to trade models that feature population mobility. They relate the relevant elasticities to reduced-form "local" and "global" elasticities that have economic meaning in the economic geography context and transform the system so that the comparative statics can be analyzed using the Neumann series representation.

Building on their work, the present paper generalizes comparative statics from trade and economic geography to all spatial models under the constant elasticity framework. Thus, the comparative statics in the present paper can be applied to a wide variety of spatial models, with an arbitrary number of interactions and notions of "flow." In this setting, we focus on the relationship between the zeroth-degree effect of the Neumann series and partial equilibrium effects, and how they can be used for economic insights and quantitative predictions. We show how alternative definitions of "partial equilibrium" effects can be applied, leading to different Neumann series with different expressions for the zeroth effect and therefore different expressions for the partial equilibrium effect that may have different relevance, both quantitatively and economically, depending on the context. Two definitions, in particular, are emphasized. One definition holds constant economic activity that is mediated through the spatial network. Loosely speaking, this means holding constant economic activity in all other locations. The other definition holds constant economic activity in other locations as well as other endogenous variables in the present location. In our quantitative application, we then apply these tools and provide additional theoretical results based on the matrix structure of our setting and apply the Neumann series to provide sufficient conditions for when population shifts within a city.

Comparative statics in spatial models have been studied in many other contexts. Miyauchi et al. (2021) derive partial equilibrium effects for a shock to the price index of non-traded services in a commuting model with trip chains. Monte et al. (2018) show the partial equilibrium effects of a productivity shock on employment and population in a model with commuting. Baqaee and Farhi (2024) derive microeconomic sufficient statistics for the general equilibrium response of variables such as output and welfare to productivity and trade cost shocks in a class of trade models with input-output networks. Kleinman et al. (2024) show that first-order comparative statics for productivity shocks in a constant elasticity economic geography model can be represented using a friend-enemy matrix that summarizes each location's exposure to productivity shocks in all locations.

The "exact-hat" technique popularized by Dekle et al. (2008) have been applied in a wide variety of models to compute exact counterfactuals using observable variables as inputs. Many of these models have been in the constant-elasticity family while others have not. We are not aware of results that articulate what data are sufficient to compute such counterfactuals for a family of models. In this respect, our results clarify what inputs are sufficient to compute exact-hat counterfactuals in the constant-elasiticy family of spatial models. We also highlight a number parallels in the exact-hat method relative to the first-order method, namely the identical data requirements and analagous partial-equilibrium expressions which can be useful for understanding comparative statics for any given model.

Our application studies the impact of commuting infrastructure on the location of residence and employment in urban areas. Previous studies have mostly emphasized the decentralizing impact of new transportation infrastructure. Glaeser and Kahn (2001) document that American cities first began to decentralize in the late 19th century as the first commuter trains and streetcars allowed workers to move their residence to the suburbs.[4] Baum-Snow (2007) use the construction of the interstate highway system as a natural experiment and find large decentralization in population from central cities. Baum-Snow et al. (2017) find large decentralizing impacts of radial and ring-road highways on population across cities in China.

The extent to which commuter rail in urban areas centralizes population and employment has been less clear cut. Heblich et al. (2020) find that the advent of the steam railway in the 19th century facilitated the large-scale shift in residential population to the suburbs and concentration of employment in the city center in absolute numbers. The authors also find similar patterns of shifting population from downtown to the suburbs in other large metropolitan areas following the transport improvements of the 19th century, including Berlin, Paris, Boston, Chicago, New York, and Philadelphia. At the same time, the highest percentage growth rates in both employment and population occurred in the suburbs. Gonzalez-Navarro and Turner (2018) found little evidence of urban population growth from subways, while Baum-Snow et al. (2017) found little evidence of population shift to the suburbs from radial railroads.

Other works on rail, such as Glaeser and Kahn (2004), Glaeser et al. (2008), and Baum-Snow et al. (2005), have documented the declining ridership and thus limited economic viability of urban rail given the high fixed costs.

Our theoretical results from our commuting model are closely related to Thisse et al. (2024), who study the equilibrium characteristics of a linear and symmetric 3-location quan-

[4]See also *Crabgrass Frontier* by Kenneth Jackson for a thorough historical account.

titative model (two peripheries and a center) and find that models with preference heterogeneity lead to an "average preference for central employment and residence" because the central location has an advantage as the place where average transportation costs are lowest. They note in particular that "when the location of production and residence is endogenous, the standard intuition that lowering commuting costs leads to dispersed economic activity need not apply."

The present paper provides a framework for analyzing the mechanisms behind commuting models using observable sufficient statistics (namely the matrix of observed commuting flows, and in particular the commuting shares for pairs of locations that experience the commuting improvement) and offers theoretical results on the tendency of radial infrastructure to promote depopulation and central-focused infrastructure to promote centralization.

From an empirical perspective, our model-based results showing that Tokyo's train system causes centralization in both population and employment are an important counterexample to the prevailing empirical evidence on the decentralizing impacts of transportation, and our theoretical results help to explain why—namely the dense subway network in the city's core.

2 Comparative Statics in Constant Elasticity Spatial Models

We consider a constant-elasticity spatial model that is widely used in spatial economics and discussed in Remark 5 of Allen et al. (2024). We follow a similar notation. Let $\mathcal{N} \equiv \{1, \ldots, N\}$ and $\mathcal{H} \equiv \{1, \ldots, H\}$ denote the set of locations and the set of economic interactions, respectively.

Let $\mathbf{x}$ be an $N \cdot H$ column vector of endogenous economic outcomes, with H blocks, each consisting of N variables:

$$\mathbf{x} \equiv \left[x_{1,1}, \ldots, x_{N,1} \,\middle|\, x_{1,2}, \ldots, x_{N,2} \,\middle|\, \cdots \,\middle|\, x_{1,H}, \ldots, x_{N,H} \right]' \in \mathbb{R}^{HN}$$

For each $i \in \mathcal{N}$ and $h \in \mathcal{H}$, x_{ih} refers to the i^{th} element in the h^{th} block of the vector $\mathbf{x}$, and is strictly positive: $\{x_{ih}\}_{i \in \mathcal{N}, h \in \mathcal{H}} \in \mathbb{R}^{NH}_{++}$. The economic system can be expressed with the following system of equations:

$$\prod_{h' \in \mathcal{H}} x_{ih'}^{\gamma_{hh'}} = \sum_{j \in \mathcal{N}} K_{ijh} \prod_{h' \in \mathcal{H}} x_{ih'}^{\rho_{hh'}} x_{jh'}^{\beta_{hh'}} \tag{1}$$

for all $i \in \mathcal{N}$ and $h \in \mathcal{H}$, where $\gamma_{hh'}, \rho_{hh'}, \beta_{hh'}$ are the (h, h')-th elements of matrices $\mathbf{\Gamma}, \mathbf{R}$, and $\mathbf{B}$, respectively. For all $i \in \mathcal{N}, j \in \mathcal{N}, h \in \mathcal{H}$, the function $K_{ijh} : \Theta \to \mathbb{R}_{++}$ is strictly positive and continuously differentiable in the parameter vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m) \in \Theta \subseteq \mathbb{R}^{M}_{++}$.[5] In our commuting model, $\boldsymbol{\theta}$ contains all exogenous fundamentals and commuting costs, $\bar{A}_i, \bar{U}_i, d_{ij}$.

For notational simplicity, we suppress the dependence on $\boldsymbol{\theta}$ and write K_{ijh} instead of $K_{ijh}(\boldsymbol{\theta})$, except when the dependence needs to be explicit.

[5] As discussed in Allen et al. (2024), it is common (especially in economic geography models) for the equilibrium system to involve an endogenous scalar λ_h for one or more interactions such that the system can be written as:

$$\prod_{h' \in \mathcal{H}} x_{ih'}^{\gamma_{hh'}} = \lambda_h \sum_{j \in \mathcal{N}} K_{ijh} \prod_{h' \in \mathcal{H}} x_{ih'}^{\rho_{hh'}} x_{jh'}^{\beta_{hh'}} \tag{2}$$

. For such models, Allen et al. (2024) show and we demonstrate in our urban application that the endogenous scalars can be absorbed into the other endogenous variables such that the system can be transformed into that of equations (1). Our comparative statics results are then regarding these transformed variables that may differ in levels from the actual endogenous variable of interest, depending on the model's assumptions on aggregate resource constraints (e.g., population mobility assumptions for the case of economy geography models). In such settings, the present formulation is immediately useful for studying comparative statics of relative changes: how one location's endogenous variables changes relative to another location's.

Rearranging (1) to isolate x_{ih}, we have:

$$x_{ih} = \left[\prod_{h' \neq h} x_{ih'}^{\rho_{hh'} - \gamma_{hh'}} \sum_{j \in \mathcal{N}} K_{ijh} \prod_{h' \in \mathcal{H}} x_{jh'}^{\beta_{hh'}} \right]^{\frac{1}{\gamma_{hh} - \rho_{hh}}} \tag{3}$$

Define $f_{ih} : \Theta \times \mathbb{R}^H_{++} \to \mathbb{R}_{++}$ as the right-hand side of equation (3):

$$f_{ih}(\mathbf{x}, \boldsymbol{\theta}) \equiv \left[\prod_{h' \neq h} x_{ih'}^{\rho_{hh'} - \gamma_{hh'}} \sum_{j \in \mathcal{N}} K_{ijh}(\boldsymbol{\theta}) \prod_{h' \in \mathcal{H}} x_{jh'}^{\beta_{hh'}} \right]^{\frac{1}{\gamma_{hh} - \rho_{hh}}} \tag{4}$$

Definition. Let $\kappa = \theta_k$ denote a generic element of the parameter vector $\boldsymbol{\theta}$.

The *single-variable local partial equilibrium effect* of a change in parameter $d \ln \kappa$ on endogenous variable x_{ih} is defined as the derivative:

$$\frac{\partial^{SLP} \ln x_{ih}}{\partial \ln \kappa} \equiv \sum_j \frac{\partial \ln f_{ih}}{\partial \ln K_{ijh}} \frac{\partial \ln K_{ijh}}{\partial \ln \kappa}.$$

One interpretation is to view the partial equilibrium effect as the effect of a change in κ on x_{ih} while holding constant all other endogenous variables in interaction h. This is not entirely correct, because x_{ih} also appears on the right-hand side of the equation. However, x_{ih} only appears there as a factor in one element of a summation, which characterizes the spatial effects. Therefore, a more accurate (though perhaps less precise) interpretation of the partial equilibrium effect is as a "local effect," holding constant spatial (i.e., cross-location) spillover forces as well as other endogenous variables located in location i, $x_{ih'}$ for $h' \neq h$.[6]

It is also useful to define a related expression and function that captures a related partial derivative. Rearranging (1) to isolate variables with i-indexed terms, we have:

$$\prod_{h' \in \mathcal{H}} x_{ih}^{\gamma_{hh'} - \rho_{hh'}} = \sum_{j \in \mathcal{N}} K_{ijh} \prod_{h' \in \mathcal{H}} x_{jh'}^{\beta_{hh'}} \tag{5}$$

Define $g_{ih} : \Theta \times \mathbb{R}^H_{++} \to \mathbb{R}_{++}$ as the right-hand side of equation (5):

$$g_{ih}(\mathbf{x}, \boldsymbol{\theta}) \equiv \sum_{j \in \mathcal{N}} K_{ijh}(\boldsymbol{\theta}) \prod_{h' \in \mathcal{H}} x_{jh'}^{\beta_{hh'}} \tag{6}$$

As in Allen et al. (2024), we define the change of variables $y_{ih} \equiv \prod_{h' \in \mathcal{H}} x_{ih}^{\gamma_{hh'} - \rho_{hh'}}$. If

[6] In the commuting model described in Section 3, the left-hand-side variable does not appear in function f_{ih} because we ruled out spillovers to neighboring locations. Thus, the partial equilibrium effect in that commuting model can be interpreted simply as "holding constant all other endogenous variables."

$\mathbf{\Gamma} - \mathbf{R}$ is invertible, we can rewrite the system and define a new function $\tilde{g}_{ih}(\mathbf{y}, \boldsymbol{\theta})$ as follows:

$$y_{ih} = \sum_{j \in \mathcal{N}} K_{ijh} \prod_{h' \in \mathcal{H}} y_{jh'}^{\alpha_{hh'}}, \qquad \tilde{g}_{ih}(\mathbf{y}, \boldsymbol{\theta}) \equiv \sum_{j \in \mathcal{N}} K_{ijh} \prod_{h' \in \mathcal{H}} y_{jh'}^{\alpha_{hh'}} \tag{7}$$

where $\alpha_{hh'}$ is the (h, h')-th element of matrix $\boldsymbol{\alpha} \equiv \mathbf{B}(\mathbf{\Gamma} - \mathbf{R})^{-1}$. We note that $g_{ih}(\cdot, \boldsymbol{\theta})$ and $\tilde{g}_{ih}(\cdot, \boldsymbol{\theta})$ are related through the change of variables:

$$\tilde{g}_{ih}(\mathbf{y}, \boldsymbol{\theta}) = g_{ih}\Big(\exp\Big(((\mathbf{\Gamma} - \mathbf{R})^{-1} \otimes \mathbf{I}_N) \ln(\mathbf{y}) \Big), \boldsymbol{\theta} \Big),$$

where $\otimes$ denotes the Kronecker product. We follow the same notation for $\mathbf{y}$ as $\mathbf{x}$:

$$\mathbf{y} \;\equiv\; \Big[y_{1,1}, \ldots, y_{N,1} \;\Big|\; y_{1,2}, \ldots, y_{N,2} \;\Big|\; \cdots \;\Big|\; y_{1,H}, \ldots, y_{N,H} \Big]' \in \mathbb{R}^{HN}$$

We refer to y_{ih} as the *local composite* variable for interaction h, because it captures all factors specific to a given location and interaction except those factors in a given location and interaction that are mediated by spatial distance: K_{ijh} when $j = i$.

Definition. The *local partial equilibrium effect* of a change in parameter $d \ln \kappa$ on endogenous variable x_{ih} is defined as the h^{th} element of the vector:

$$\frac{\partial^{LP} \ln(\mathbf{x}_i)}{\partial \ln \kappa} \equiv (\mathbf{\Gamma} - \mathbf{R})^{-1} \frac{\partial^{CLP} \ln(\mathbf{y}_i)}{\partial \ln \kappa} = (\mathbf{\Gamma} - \mathbf{R})^{-1} \frac{\partial \ln \mathbf{g}_i}{\partial \ln \kappa},$$

where $\mathbf{y}_i$ and $\mathbf{x}_i$ are vectors whose h-th elements are y_{ih} and x_{ih} respectively, and $\frac{\partial^{CLP} \ln(\mathbf{y}_i)}{\partial \ln \kappa} = \frac{\partial \ln \mathbf{g}_i}{\partial \ln \kappa}$ is a vector whose h-th element is

$$\sum_j \frac{\partial \ln g_{ih}}{\partial \ln K_{ijh}} \frac{\partial \ln K_{ijh}}{\partial \ln \kappa}.$$

To describe these variables in matrix notation, define the composite local partial equilibrium effect vector:

$$b_{ih} := \sum_j \frac{\partial \ln g_{ih}}{\partial \ln K_{ijh}} \frac{\partial \ln K_{ijh}}{\partial \ln \kappa}, \qquad \mathbf{b} = \Big[b_{1,1}, \ldots, b_{N,1} \;\Big|\; \cdots \;\Big|\; b_{1,H}, \ldots, b_{N,H} \Big]'.$$

Then the vectorized partial equilibrium effect is given by

$$\frac{\partial^{LP} \ln(\mathbf{x})}{\partial \ln \kappa} = \Big((\mathbf{\Gamma} - \mathbf{R})^{-1} \otimes I_N \Big) \mathbf{b},$$

which is a column vector. If we wish to consider the entire parameter vector $\boldsymbol{\theta}$, we replace

κ with the row vector $\boldsymbol{\theta}'$.

The purpose of the composite variable is to isolate variables that are specific to a location for a particular interaction. Thus, the right-hand side of $g_{ih}(\cdot)$ excludes other local variables except when mediated by the spatial factor K_{ijh} when $j = i$. Therefore, the partial equilibrium impact on the composite variable holds constant the composite variables in other locations but allows the other composite variables in the same location to change. This interpretation contrasts with that of the *single-variable* local partial equilibrium effect, where all other endogenous variables in the current location are held constant.

The local partial equilibrium effect of a change in parameter $d\ln\kappa$ on endogenous variable x_{ih} thus incorporates local equilibrium forces from multiple interactions and summarizes the full effect of these forces while holding constant spatial spillover forces.

We are now ready to state our theorems.

Theorem 1 (Comparative Statics)**.**
Part 1. *Let* $\boldsymbol{\gamma} \equiv (\boldsymbol{\Gamma} - \mathbf{R})$ *and assume it is non-singular so that* $\boldsymbol{\alpha} = \mathbf{B}\,\boldsymbol{\gamma}^{-1}$*. Denote a solution to the system in* (1) *as a stacked column vector*

$$\mathbf{x}^* \equiv \left[\, x^*_{1,1}, \ldots, x^*_{N,1} \;\middle|\; x^*_{1,2}, \ldots, x^*_{N,2} \;\middle|\; \cdots \;\middle|\; x^*_{1,H}, \ldots, x^*_{N,H} \,\right]' \in \mathbb{R}^{HN},$$

with parameters $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_M]'$ *as before.*

Then, given a solution and whenever $\mathbf{A}$ *is nonsingular, the elasticities of endogenous variables with respect to parameters* $\boldsymbol{\theta} \in \mathbb{R}^M$ *are*

$$\frac{\partial \ln \mathbf{x}^*}{\partial \ln \boldsymbol{\theta}'} = -\mathbf{A}^{-1}\mathbf{T}, \tag{8}$$

where $\mathbf{A} \equiv \overline{\mathbf{A}}\,\tilde{\boldsymbol{\Gamma}}$ *and*

$$\tilde{\boldsymbol{\Gamma}} \equiv \boldsymbol{\gamma} \otimes \mathbf{I}_N, \qquad \overline{\mathbf{A}} \equiv \mathbf{I}_{HN} - \boldsymbol{\alpha}_X, \qquad \boldsymbol{\alpha}_X \equiv \operatorname{bdiag}\Big(\mathbf{X}_{(1)}, \ldots, \mathbf{X}_{(H)}\Big)(\boldsymbol{\alpha} \otimes \mathbf{I}_N),$$

where $\mathbf{I}$ *is an identity matrix, and* $\otimes$ *denotes the Kronecker product.*

For each $h \in \{1, \ldots, H\}$*, let the "flow share" matrix* $\mathbf{X}_{(h)} \in \mathbb{R}^{N\times N}$ *be the row-stochastic matrix (each row sums to one) with entries:*

$$\Big[\mathbf{X}_{(h)}\Big]_{ij} = \frac{K_{ijh}\,\prod_{h'\in\mathcal{H}} y_{jh'}^{\alpha_{hh'}}}{\sum_{j'} K_{ij'h}\,\prod_{h'\in\mathcal{H}} y_{j'h'}^{\alpha_{hh'}}} = \frac{K_{ijh}\,\prod_{h'\in\mathcal{H}} x_{jh'}^{\beta_{hh'}}}{\prod_{h'\in\mathcal{H}} x_{ih'}^{\gamma_{hh'}-\rho_{hh'}}}. \tag{9}$$

Here we have used the block–diagonal operator

$$\mathrm{bdiag}\Big(\mathbf{X}_{(1)},\ldots,\mathbf{X}_{(H)}\Big) \equiv \begin{bmatrix} \mathbf{X}_{(1)} & 0 & \cdots & 0 \\ 0 & \mathbf{X}_{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{X}_{(H)} \end{bmatrix} \in \mathbb{R}^{HN\times HN}.$$

Last, $\mathbf{T} \in \mathbb{R}^{HN\times M}$ *is block-stacked as*

$$\mathbf{T} = \Big[\mathbf{T}_1' \ \cdots \ \mathbf{T}_H'\Big]',$$

with $\mathbf{T}_h \in \mathbb{R}^{N\times M}$ *having entries:*

$$\Big[\mathbf{T}_h\Big]_i = \sum_{j=1}^{N} \Big[\mathbf{X}_{(h)}\Big]_{ij} \frac{\partial \ln K_{ijh}}{\partial \ln \boldsymbol{\theta}'} \qquad (h = 1,\ldots,H;\ i = 1,\ldots,N). \tag{10}$$

Part 2. *Assume* $\boldsymbol{\Gamma}$ *is nonsingular. Then* $\mathbf{A}^{-1}$ *can be expressed as the following Neumann series if and only if* $\rho(\boldsymbol{\alpha}_X) < 1$*:*

$$\mathbf{A}^{-1} = \tilde{\boldsymbol{\Gamma}}^{-1} \sum_{k=0}^{\infty} \boldsymbol{\alpha}_X^k. \tag{11}$$

The ∞ *matrix norm is given by*

$$\|\boldsymbol{\alpha}_X\|_\infty = \| \ |\boldsymbol{\alpha}_X| \ \|_\infty = \|\boldsymbol{\alpha}\|_\infty = \| \ |\boldsymbol{\alpha}| \ \|_\infty.$$

Thus, the terms $\boldsymbol{\alpha}_X^k$ *decay according to*

$$\|\boldsymbol{\alpha}_X^k\|_\infty \leq \| \ |\boldsymbol{\alpha}| \ \|_\infty^k \qquad \textit{for all } k.$$

Moreover, $\rho(|\boldsymbol{\alpha}_X|) = \rho(|\boldsymbol{\alpha}|)$*, while* $\rho(\boldsymbol{\alpha}) \leq \rho(\boldsymbol{\alpha}_X) \leq \rho(|\boldsymbol{\alpha}_X|)$*. Thus,* $\rho(|\boldsymbol{\alpha}|) < 1$ *guarantees the Neumann representation.*

(ii) Define

$$\overline{D}_{\boldsymbol{\theta}}^k \ln \mathbf{x}^* \equiv -\tilde{\boldsymbol{\Gamma}}^{-1} \boldsymbol{\alpha}_X^k \mathbf{T}$$

as the k^{th}*-degree term of the comparative static (i.e., such that* $\frac{\partial \ln \mathbf{x}^*}{\partial \ln \boldsymbol{\theta}'} = \sum_{k=0}^{\infty} \overline{D}_{\boldsymbol{\theta}}^k$*).*

The zeroth-degree term of the comparative static is equal to the local partial equilibrium effect from a change in parameters $d\ln\boldsymbol{\theta}$ *on variables* $\mathbf{x}$*:*

$$\overline{D}_{\boldsymbol{\theta}}^0 \ln \mathbf{x}^* = -\tilde{\boldsymbol{\Gamma}}^{-1}\mathbf{T} = \frac{\partial^{LP} \ln(\mathbf{x})}{\partial \ln \boldsymbol{\theta}'}.$$

(iii) Define

$$\mathbf{R} = \tilde{\mathbf{\Gamma}}^{-1} \sum_{k=1}^{\infty} \boldsymbol{\alpha}_X^k,$$

and the general equilibrium remainder of the comparative static as

$$\sum_{k=1}^{\infty} \overline{D}_{\boldsymbol{\theta}}^k = -\mathbf{RT},$$

so that

$$-\frac{\partial \ln \mathbf{x}^*}{\partial \ln \boldsymbol{\theta}'} = \mathbf{RT} + \tilde{\mathbf{\Gamma}}^{-1}\mathbf{T}.$$

Now, consider a series of comparative statics:

$$d \ln x_{lh} = \sum_m \frac{\partial \ln x_{lh}}{\partial \ln \theta_m} \, d \ln \theta_m.$$

Let $\mathbf{e}_{lh} \in \mathbb{R}^{NH}$ *denote the vector with 1 in the* l^{th} *element of the* h^{th} *block and 0 in all others. Let* $\Delta \ln \boldsymbol{\theta} \in \mathbb{R}^M$ *be the vector of log changes to parameters* $\boldsymbol{\theta}$ *such that* $[\Delta \ln \boldsymbol{\theta}]_m = \Delta \ln \theta_m$*. Therefore,*

$$\Delta \ln x_{lh} = \mathbf{e}_{lh}' \frac{\partial \ln \mathbf{x}^*}{\partial \ln \boldsymbol{\theta}'} \Delta \ln \boldsymbol{\theta}.$$

Then two bounds for the general equilibrium remainder of the comparative static are as follows:

$$|\mathbf{e}_{lh}' \mathbf{RT} \Delta \ln \boldsymbol{\theta}| \leq \|\Delta \ln \boldsymbol{\theta}\|_\infty \, \|\tilde{\mathbf{\Gamma}}^{-1}\|_\infty \, \|\mathbf{T}\|_\infty \frac{\|\boldsymbol{\alpha}_X\|_\infty}{1 - \|\boldsymbol{\alpha}_X\|_\infty}$$
$$|\mathbf{e}_{lh}' \mathbf{RT} \Delta \ln \boldsymbol{\theta}| \leq \|\Delta \ln \boldsymbol{\theta}\|_\infty \left\| \left(\mathbf{e}_{lh} \tilde{\mathbf{\Gamma}}^{-1} \right)' \right\|_1 \|\mathbf{T}\|_\infty \frac{\|\boldsymbol{\alpha}_X\|_\infty}{1 - \|\boldsymbol{\alpha}_X\|_\infty}$$

Similarly, when considering relative changes between two endogenous variables x_{lh} *and* $x_{l'h'}$*:*

$$d \ln \left(\frac{x_{lh}}{x_{l'h'}} \right) = \sum_m \left[\frac{\partial \ln x_{lh}}{\partial \ln \theta_m} - \frac{\partial \ln x_{l'h'}}{\partial \ln \theta_m} \right] d \ln \theta_m,$$

let $\mathbf{u} \in \mathbb{R}^{NH}$ *denote the vector with* 1 *in the* l^{th} *element of the* h^{th} *block,* -1 *in the* l'^{th} *element of the* h'^{th} *block, and 0 elsewhere.*

Then two bounds for the general equilibrium remainder of this expression are as follows:

$$|\mathbf{u}'\mathbf{R}\mathbf{T}\Delta\ln\boldsymbol{\theta}| \leq 2\|\Delta\ln\boldsymbol{\theta}\|_\infty \|\tilde{\boldsymbol{\Gamma}}^{-1}\|_\infty \|\mathbf{T}\|_\infty \frac{\|\boldsymbol{\alpha}_X\|_\infty}{1-\|\boldsymbol{\alpha}_X\|_\infty}$$
$$|\mathbf{u}'\mathbf{R}\mathbf{T}\Delta\ln\boldsymbol{\theta}| \leq \|\Delta\ln\boldsymbol{\theta}\|_\infty \left\|\left(\mathbf{u}'\tilde{\boldsymbol{\Gamma}}^{-1}\right)'\right\|_1 \|\mathbf{T}\|_\infty \frac{\|\boldsymbol{\alpha}_X\|_\infty}{1-\|\boldsymbol{\alpha}_X\|_\infty}$$

A number of comments are in order.

First, the theorem provides a toolkit for conducting comparative-static analysis across a wide range of spatial models. In many situations, the flow-share matrix $\mathbf{X}$ represents a meaningful and often observable measure of flows between locations. In our commuting model, $\mathbf{X}$ consists of two interactions: residence and employment. For the residence interaction, $\mathbf{X}_1 = \frac{\mathbf{L}}{\mathbf{L_R}}$ is the share of a location's residents who commute *to* a workplace, while for the employment interaction, $\mathbf{X}_2 = \frac{\mathbf{L}}{\mathbf{L_F}}$ is the share of a workplace's commuters who commute *from* a location.

In many gravity-based trade models, one interaction of $\mathbf{X}$ is the share of a recipient's expenditure coming from an exporter, while another is the share of production value sent to a recipient (see, e.g., Allen et al. (2020)). Thus, in many cases, knowledge of the fundamentals contained in K_{ijh} is not required for comparative statics. In others, the parametrization of K_{ijh} (together with observable location-specific data) is sufficient, as in our empirical application.

Given the matrix of "flow shares" for each interaction, and elasticities of the model, one can therefore calculate the matrix of first-order comparative static elasticities, $\frac{\partial \ln \mathbf{x}^*}{\partial \ln \boldsymbol{\theta}'}$. If the researcher also knows the change in parameters $d\ln\boldsymbol{\theta}$, they can then compute the first-order change to endogenous variables. In many cases, the matrix of *flows* (a variable proportional to $K_{ijh}\prod_{h'\in\mathcal{H}} x_{jh'}^{\beta_{hh'}}$) is more readily available than the matrix of *flow shares*. In such a situation, the researcher can construct the flow-share matrix by dividing the flow by the sum of the flows for a particular location. By summing the flows for a location, the researcher can also obtain the values of endogenous variables for the current equilibrium. The researcher then not only knows the first-order changes to equilibrium quantities but also the initial quanitities and therefore have the requirements to compute a counterfactual equilibrium, rather than simply *changes* to endogenous variables.

Relatedly, we note that the data that is sufficient to conduct comparative statics (i.e. data on flow shares) is invariant to the researcher's choice of endogenous variables. For many systems of equations, the researcher makes a choice regarding the endogenous variable of interest depending on the empirical context. The same system of equations can lead to multiple formulations with different endogenous variables. So long as the mapping for a new change of variables is invertible (at least locally), positive, and continuously differentiable, the implicit function can be applied and the same data on flow-shares are sufficient to compute comparative statics.[7] For example, consider a common scenario where endogenous variables are related in a log-linear fashion, and we have computed the comparative static for x_{hi} $i = 1, \ldots, N; h = 1, \ldots, H$, and we seek instead to obtain the comparative static equations for $\tilde{x}_{ih}$ which are related to x_{ih} by the following function that is loglinear to an exogenous location-specific constant and other endogenous variables: $\tilde{x}_{ih} = D_{ih} \prod_{h'} x_{h'i}^{\eta_{hh'}}$ for $i = 1, \ldots, N; h = 1, \ldots, H$, where $\eta_{hh'}$ is (h, h') the element of a matrix $\boldsymbol{\eta}$, and D_{ih} is exogenous. Denoting the $\bar{\mathbf{x}}_i$ to be a vector whose elements are equal to $\bar{x}_i h = \tilde{x}_{ih}/D_{ih}$, we have the matrix relation, $\ln(\bar{\mathbf{x}}_i) = \boldsymbol{\eta} \ln(\mathbf{x}_i)$, and if $\boldsymbol{\eta}$ is invertible, $\ln(\mathbf{x}_i) = \boldsymbol{\eta}^{-1} \ln(\bar{\mathbf{x}}_i)$. Moreover, the relationship in changes is independent of the location-specific variable: $d\ln(\bar{\mathbf{x}}_i) = d\ln(\tilde{\mathbf{x}}_i) = \boldsymbol{\eta} d\ln(\mathbf{x}_i)$. When the matrix $\boldsymbol{\eta}$ is invertible, then the implicit function theorem applies to the transformed function, and we can obtain the new comparative static using a transformation of the previous one: $\frac{\partial \ln \tilde{\mathbf{x}}^*}{\partial \ln \boldsymbol{\theta}'} = (\boldsymbol{\eta} \otimes \mathbf{I}) \frac{\partial \ln \mathbf{x}^*}{\partial \ln \boldsymbol{\theta}'}$. Thus both comparative statics can be computed using data on flow-shares. While these choices may lead to new elasticities to consider (e.g. the matrix $\boldsymbol{\eta}$), they may still utilize the same flow-share data.

Second, consider some well-known parameterizations of K_{ijh}. If θ_k depends on both locations i and j, as in gravity models ($\theta_k = d_{i'j'h}$) with $K_{ijh} \propto d_{ijh}^{\phi}$, then

$$\Big[\mathbf{T}_h\Big]_{ik} = \phi \Big[\mathbf{X}_{(h)}\Big]_{ij'} \quad \text{if } i = i', \quad 0 \text{ otherwise.}$$

This describes the comparative static for location i in interaction h. In gravity models with sender and receiver interactions, it is often the case that $d_{i'j'h}$ appears in both interactions.

Suppose we have a transport cost matrix $\boldsymbol{D}$, with $d_{\cdot,l}$ denoting the l^{th} column, $d_{l,\cdot}$ the l^{th} row, and $d_{l,k}$ the element in row l and column k. For a separate interaction h', we may have $d_{i'j'h} = d_{i'j'h'} = d_{i'j'}$ for $i', j' = 1, \ldots, N$ and $K_{ijh'} \propto d_{jih'}^{\phi} = d_{ji}^{\phi}$, implying

$$\Big[\mathbf{T}_{h'}\Big]_{ik} = \phi \Big[\mathbf{X}_{(h')}\Big]_{ij'} \quad \text{if } i = i', \quad 0 \text{ otherwise.}$$

Because $K_{ijh} \propto d_{ij}^{\phi}$ in one interaction and $K_{ijh} \propto d_{ji}^{\phi}$ in the other, the two interactions differ

[7] For example, in the international-trade context, Allen et al. (2020) use a trade model of the constant-elasticity type (e.g. (Arkolakis et al. (2012))) and choose export and import prices as their endogenous variables of interest to understand the impact of a hypothetical trade war with China. In their results using a constant-elasticity trade model, Kleinman et al. (2024) use incomes and welfare to study the impact of productivity growth in one country on the income and welfare of other countries. Underlying these two models are the same balanced-trade assumptions from Arkolakis et al. (2012). These two choices of endogenous variables are related via a log-linear transformation, and therefore flow-share data can be used in both contexts.

structurally when stacked in matrix form.

For example, let $\boldsymbol{\theta} = \mathrm{vec}(\boldsymbol{D})$ (stacking columns of $\boldsymbol{D}$). If d_{ij} denotes the cost from sender (e.g., exporter or residence) to receiver (e.g., importer or workplace), denote the sending interaction as $h = s$ and the receiving interaction as $h = r$ (with some abuse of notation). Then:

$$\left[\mathbf{T}_s\right] = \phi\Big(\mathbf{X}_{(s)} \otimes \mathbf{1}\Big) \circ \Big(\mathbf{1} \otimes \mathbf{I}\Big), \qquad \left[\mathbf{T}_r\right] = \phi\Big(\mathbf{1} \otimes \mathbf{X}_{(r)}\Big) \circ \Big(\mathbf{I} \otimes \mathbf{1}\Big).$$

In trade models, $\mathbf{X}_{(s)}$ is the export-share matrix (share of each country's exports going to each other country), while $\mathbf{X}_{(r)}$ is the import-share matrix. The resulting comparative static is then:[8]

$$\frac{\partial \ln x_{sl}}{\partial \ln d_{ij}} = [\mathbf{A}^{-1}_{s,s}]_{li}\, \phi[\mathbf{X}_{(r)}]_{ij} + [\mathbf{A}^{-1}_{s,r}]_{l,j}\, \phi[\mathbf{X}_{(s)}]_{ji} + \dots, \qquad l, i, j = 1, \dots, N, \tag{12}$$

$$\frac{\partial \ln x_{rl}}{\partial \ln d_{ij}} = [\mathbf{A}^{-1}_{r,s}]_{l,i}\, \phi[\mathbf{X}_{(s)}]_{ij} + [\mathbf{A}^{-1}_{r,r}]_{l,j}\, \phi[\mathbf{X}_{(r)}]_{ji} + \dots, \qquad l, i, j = 1, \dots, N. \tag{13}$$

Here $[\mathbf{A}^{-1}_{h,h'}]_{ij}$ denotes the (i, j) element of the $(h, h')^{\text{th}}$ block of $\mathbf{A}^{-1}$, and the $\dots$ indicate additional terms if other interactions include d_{ij} in the equilibrium equation.

In trade models with goods-market clearing and balanced trade, x_{sl} may represent exporter prices and x_{rl} importer price indices (see Allen et al. (2020)). In other contexts, endogenous variables may lack a clear sender/receiver interpretation. For example, in economic geography, Allen and Arkolakis (2025) use wages w_i and population L_i. Despite different definitions, the structure of (12) remains: a weighted sum of terms, with weights given by observable flow-share matrices $[\mathbf{X}_{(s)}]$ and $[\mathbf{X}_{(r)}]$.[9]

If θ_k depends on a left-hand-side–indexed parameter, e.g. $\theta_k = A_k$ and $K_{ijh} \propto A_i$ for some h, then

$$[\mathbf{T}_h]_{ik} = 1 \quad \text{if } i = k, \quad 0 \text{ otherwise},$$

so $\mathbf{T}_{(h)} = \mathbf{I}$.

If θ_k depends on a right-hand-side–indexed parameter, e.g. $\theta_k = A_k$ and $K_{ijh} \propto A_j$, then $[\mathbf{T}_h]_{ik} = [\mathbf{X}_{(h)}]_{ik}$.

In a gravity setting where exogenous fundamentals appear in both interaction equations—for instance $K_{ijh} \propto U_i A_j$ for interaction h and $K_{ijh'} \propto A_i U_j$ for h'—we stack the

[8] See Allen et al. (2020) for an alternative derivation.

[9] This structure can change if one applies a monotonic transformation to both sides of (1) or (7), so that the right-hand side is not a simple sum across locations. In that case, the comparative statics remain the same, but $\mathbf{T}$ and $\mathbf{A}$ take a different form; see Allen et al. (2020).

identity and flow-share matrices. With $\boldsymbol{\theta} = [\mathbf{U}, \mathbf{A}]' \in \mathbb{R}^{2N}$ and $[\mathbf{T}_h, \mathbf{T}_{h'}] \in \mathbb{R}^{2N\times 2N}$, we have:

$$\begin{bmatrix} \mathbf{T}_h \\ \mathbf{T}_{h'} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_N & \mathbf{X}_{(h)} \\ \mathbf{X}_{(h')} & \mathbf{I}_N \end{bmatrix}.$$

The element-wise comparative statics are given by:

$$\frac{\partial \ln x_{hl}}{\partial \ln U_i} = [\mathbf{A}^{-1}_{h,h}]_{ll}\,\delta_{l,i} + \sum_{j=1}^{N} [\mathbf{A}^{-1}_{h,h'}]_{l,j} [\mathbf{X}_{(h)}]_{i,j} + \cdots, \qquad l, i = 1, \ldots, N, \tag{14}$$

$$\frac{\partial \ln x_{hl}}{\partial \ln A_i} = \sum_{j=1}^{N} [\mathbf{A}^{-1}_{h,h}]_{l,j} [\mathbf{X}_{(h)}]_{j,i} + [\mathbf{A}^{-1}_{h,h'}]_{ll}\,\delta_{l,i} + \cdots, \qquad l, i = 1, \ldots, N, \tag{15}$$

$$\frac{\partial \ln x_{h'l}}{\partial \ln U_i} = \sum_{j=1}^{N} [\mathbf{A}^{-1}_{h',h'}]_{l,j} [\mathbf{X}_{(h')}]_{j,i} + [\mathbf{A}^{-1}_{h',h}]_{ll}\,\delta_{l,i} + \cdots, \qquad l, i = 1, \ldots, N, \tag{16}$$

$$\frac{\partial \ln x_{h'l}}{\partial \ln A_i} = [\mathbf{A}^{-1}_{h',h'}]_{ll}\,\delta_{l,i} + \sum_{j=1}^{N} [\mathbf{A}^{-1}_{h',h}]_{l,j} [\mathbf{X}_{(h)}]_{j,i} + \cdots, \qquad l, i = 1, \ldots, N, \tag{17}$$

Here, $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.

Third, Part 2 establishes conditions that guarantee a formulation of the comparative statics which makes explicit that the partial-equilibrium effect is the largest term of a decaying Neumann series. It also provides upper bounds on the size of the remainder term for an arbitrary series of parameter changes. This allows a researcher to make predictions and understand mechanisms about the full general equilibrium effect without explicitly calculating the entire general equilibrium. Using Part (iii), if we conduct a comparative static by increasing all parameters by 1%, the general equilibrium remainder for the change to one of the endogenous variables corresponding to interaction h is less than

$$\|\boldsymbol{\gamma}_h^{-1}\|_\infty \, \|\mathbf{T}\|_\infty \, \frac{\|\boldsymbol{\alpha}_X\|_\infty}{1 - \|\boldsymbol{\alpha}_X\|_\infty}$$

in magnitude where $\boldsymbol{\gamma}_h^{-1}$ corresponds to the h^{th} row of $\boldsymbol{\gamma}^{-1}$. Often, $\|\mathbf{T}\|_\infty$ can be bounded or expressed readily if we restrict parameters to a certain type—for example, only transport costs or only productivity. For instance, if $K_{ijh} \propto d_{ij}$ and we consider a 1% comparative static change of all d_{ij}, then $\|\mathbf{T}\|_\infty \leq 1$, with equality attained when all d_{ij} are included. When equality is attained, the remainder term is bounded by

$$\|\boldsymbol{\gamma}_h^{-1}\|_\infty \, \frac{\|\boldsymbol{\alpha}_X\|_\infty}{1 - \|\boldsymbol{\alpha}_X\|_\infty},$$

which whenever $\boldsymbol{\gamma}_h^{-1}$ is nonnegative, $\frac{\|\boldsymbol{\alpha}_X\|_\infty}{1-\|\boldsymbol{\alpha}_X\|_\infty}$ times the partial-equilibrium effect.[10] Analogous bounds are obtained if one changes all (or a subset of) fundamental productivities or all (or a subset of) fundamental amenities. Loosely speaking then $\frac{\|\boldsymbol{\alpha}_X\|_\infty}{1-\|\boldsymbol{\alpha}_X\|_\infty}$ is a measure of the magnitude of the general equilibrium remainder, relative to the partial equilibrium effect.

For other comparative statics that do not involve changing all parameters, the researcher can restrict attention to only the relevant columns of $\mathbf{T}$, thus shrinking the bounds further. These bounds represent the maximum possible error if one were to use the partial-equilibrium effect to approximate general-equilibrium comparative statics. The theorem therefore quantitatively establishes the importance of the partial-equilibrium effect relative to the general-equilibrium effect, and proposes a tool to exploit this property.

Fourth, an elegant aspect of this result is the parallel between the conditions guaranteeing uniqueness (and iterative computation) and those guaranteeing the Neumann representation. The condition $\rho(|\boldsymbol{\alpha}|) < 1$, which guarantees existence of a unique solution, is equivalent to $\rho(|\boldsymbol{\alpha}_X|) < 1$, which guarantees that the comparative static can be expressed as a Neumann series with the partial-equilibrium effect as its leading term. In Section B.3, we further show that whenever $\boldsymbol{\alpha}$ is sign-similar to a nonnegative matrix, $\rho(|\boldsymbol{\alpha}|) = \rho(\boldsymbol{\alpha})$. In this case, the Neumann series representation is possible if and only if the model has a unique solution for all geographies, and the rate of convergence of iterative computation of the equilibrium coincides with the rate of decay of the Neumann series.

An intuition for this parallel can be seen by analyzing the role of $|\boldsymbol{\alpha}|$ and $\rho(|\boldsymbol{\alpha}_X|)$ in each theorem. Allen et al. (2024) show that $\rho(|\boldsymbol{\alpha}|) < 1$ guarantees that iterating on the transformation function $\tilde{g}_{ih}(y, \theta)$ in (7) leads to a contraction regardless of inputs. For each interaction (h, h'), $|\alpha_{hh'}|$ in a constant-elasticity model describes the percent change in y_{ih} from a 1% change in spatial variables $y_{jh'}$ across all locations. Loosely speaking, $\rho(|\boldsymbol{\alpha}|) < 1$ ensures that these effects are "small enough" for contraction. Similarly, when spatial spillovers are sufficiently small, the general-equilibrium remainder decays, allowing a Neumann series representation in which the partial-equilibrium effect is the first and largest term, with subsequent terms representing dampened network-propagation effects mediated by $\boldsymbol{\alpha}_X$. We thus refer to $boldsymbol\alpha_X$ as a spillover matrix.

Another parallel lies in the role of variable transformations. Allen et al. (2024) show that transforming the system into the form of (7) yields sufficiency: $\rho(|\boldsymbol{\alpha}|) < 1$ implies a unique, iteratively computable solution. However, alternative changes of variables can yield different spectral radii and therefore different convergence rates (so long as the radius remains below one). In the comparative-statics context, a change of variables alters the definition of

[10]This is an artificial counterfactual. Note that if all d_{ij} change by the same amount, there would be no change to the *relative* values of the endogenous variables.

what is held constant when computing partial elasticities—that is, it changes the definition of the "partial-equilibrium effect." We show that an alternative change of variables leads to a different Neumann representation, with the zeroth-degree term again corresponding to a partial-equilibrium effect, but defined under different local hold-constant conditions, and with a potentially different decay rate. This flexibility is useful for two reasons. First, different definitions of the "partial-equilibrium effect" may be relevant in different empirical settings. Second, certain transformations can make the partial-equilibrium effect more interpretable—for example, if $\tilde{\mathbf{\Gamma}}$ is diagonal, then $\tilde{\mathbf{\Gamma}}^{-1}$ is also diagonal, yielding simple closed-form expressions.

We first redefine the partial-equilibrium effect flexibly based on what local equilibrium forces are held constant:

Definition. Let $\tilde{f}_{ih} : \mathbb{R}^{H\times H} \times \mathbb{R}^{H}_{++} \times \Theta \to \mathbb{R}^{NH}_{++}$ be defined as:

$$\tilde{f}_{ih}(\mathbf{V}, \mathbf{x}, \boldsymbol{\theta}) \equiv \prod_{h'\in\mathcal{H}} x_{ih'}^{V_{hh'}} \sum_{j\in\mathcal{N}} K_{ijh}(\boldsymbol{\theta}) \prod_{h'\in\mathcal{H}} x_{jh'}^{\beta_{hh'}}.$$

The *local partial equilibrium effect on* x_{ih}*, holding constant local equilibrium effects* $x_{ih'}^{V_{hh'}}$ *for* $h = 1,\ldots,H;\ h' = 1,\ldots,H$ from a change in parameter $d\ln\kappa$ on endogenous variable x_{ih} is defined as the h^{th} element of the vector:

$$\frac{\partial^{LP}_{\mathbf{V}} \ln(\mathbf{x}_i)}{\partial \ln\kappa} \equiv (\mathbf{\Gamma} + \mathbf{V} - \mathbf{R})^{-1} \frac{\partial \ln \tilde{\mathbf{f}}_i}{\partial \ln\kappa},$$

where $\mathbf{x}_i$ are vectors whose h-th element is x_{ih} respectively, and $\frac{\partial \ln \tilde{\mathbf{f}}_i}{\partial \ln\kappa}$ is a vector whose h-th element is

$$\sum_j \frac{\partial \ln \tilde{f}_{ih}}{\partial \ln K_{ijh}} \frac{\partial \ln K_{ijh}}{\partial \ln\kappa}.$$

This definition of the partial equilibrium effect allows one to flexibly choose what local equilibrium forces are held constant. One can move the dependence of an entire local variable (or multiple local variables) to the right-hand-side such that entire dependence on that local variable (or multiple local variables) are held constant. The single-variable local-equibrium effect is an example of this where we move the variables other than the one in question to the to the right-hand-side, thereby holding constant all local equilibrium variables other than the one in question.

One could also move *part* of the dependence of one variable, such that the remaining part is left on the left-hand side. This may be valuable if one part of the dependence may represent one equilibrium force (e.g. demand) and another part represents another force (e.g. supply) and therefore separating these equilibrium forces allows the researcher to isolate each equilibrium force and analyze them separately.[11]

Theorem 2 (Alternative Neumann Series for a Change of Variable)**.**
Part 1. *(i.) Let $\breve{\boldsymbol{\gamma}}(\mathbf{V}) \equiv \boldsymbol{\gamma} + \mathbf{V} = \boldsymbol{\Gamma} + \mathbf{V} - \mathbf{R}$ for $\mathbf{V} \in \mathbb{R}^{H\times H}$. Assume $\breve{\boldsymbol{\gamma}}(\mathbf{V})$ is invertible. Then, for a given $\mathbf{V}$, $\mathbf{A}^{-1}$ can be expressed as the following Neumann series if and only if $\rho(\breve{\boldsymbol{\alpha}}_X) < 1$:*

$$\mathbf{A}^{-1} = \breve{\boldsymbol{\Gamma}}^{-1} \sum_{k=0}^{\infty} \breve{\boldsymbol{\alpha}}_X^k,$$

where

$$\breve{\boldsymbol{\Gamma}} \equiv \breve{\boldsymbol{\gamma}}(\mathbf{V}) \otimes \mathbf{I}_N, \qquad \breve{\boldsymbol{\alpha}}_X(\mathbf{V}) \equiv (\mathbf{V}\breve{\boldsymbol{\gamma}}^{-1} \otimes \mathbf{I}_N) + \mathrm{bdiag}\Big(\mathbf{X}_{(1)}, \ldots, \mathbf{X}_{(H)}\Big)(\mathbf{B}\breve{\boldsymbol{\gamma}}^{-1} \otimes \mathbf{I}_N).$$

The terms decay according to

$$\|\breve{\boldsymbol{\alpha}}_X\|_\infty = \|\mathbf{V}\,\breve{\boldsymbol{\gamma}}^{-1}\|_\infty + \|\mathbf{B}\,\breve{\boldsymbol{\gamma}}^{-1}\|_\infty.$$

(ii.) Define

$$\overline{D}^k_{\mathbf{V}\boldsymbol{\theta}} \ln \mathbf{x}^* \equiv -\breve{\boldsymbol{\Gamma}}^{-1}\breve{\boldsymbol{\alpha}}_X^k \mathbf{T}$$

as the k^{th}-degree term of the comparative static associated with the matrix $\mathbf{V}$ (i.e., such that $\frac{\partial \ln \mathbf{x}^*}{\partial \ln \boldsymbol{\theta}'} = \sum_{k=0}^{\infty} \overline{D}^k_{\mathbf{V}\boldsymbol{\theta}}$).

The zeroth-degree term of the comparative static is equal to the local partial equilibrium effect on x_{ih}, holding constant local equilibrium variables x_{ih} for $h = 1, \ldots, H$; $, h' = 1, \ldots, H$:

$$\overline{D}^0_{\mathbf{V}\boldsymbol{\theta}} \ln \mathbf{x}^* = -\breve{\boldsymbol{\Gamma}}^{-1}\mathbf{T} = \frac{\partial^{LP}_{\mathbf{V}} \ln(\mathbf{x})}{\partial \ln \boldsymbol{\theta}'}$$

Part 2. *Choose $\mathbf{V} = \boldsymbol{\gamma} - \mathrm{diag}(\boldsymbol{\gamma})$, so that $\breve{\boldsymbol{\gamma}}(\mathbf{V}) = \mathrm{diag}(\boldsymbol{\gamma})$. Then the zeroth-degree term of the comparative static is equal to the single-variable local partial-equilibrium effect*

$$-\left[\breve{\boldsymbol{\Gamma}}^{-1}\mathbf{T}\right]_{ih} = \frac{\partial^{SLP} \ln x_{ih}}{\partial \ln \kappa}.$$

Proof. See Section 2 □

Part 1 of the theorem shows that there is flexibility in one's choice of definition of partial equlibrium effect, and that one's choice determines how to structure one's Neumann series such that the partial equilibrium effect is equal to the zeroth degree term. In Allen et al. (2020), the authors defined the partial equilibrium such that demand factors are held constant. In this case, $\breve{\boldsymbol{\Gamma}}^{-1}$ represents the change in demand, tracing along the supply curve. The

[11] See for example Allen et al. (2020).

second part of the theorem shows that one convenient definition of partial equilibrium (the single-variable local partial equilibrium effect) corresponds to the zeroth degree comparative static for a particular choice of $\mathbf{V}$.

Both local and single-variable local effects can potentially be incorporated into comparative-static analysis, but they have different requirements and interpretations. Incorporating the single-variable effect has the benefit of interpretability: it is simply the partial elasticity of the right-hand side of one equilibrium equation (3), holding all other variables and spillover effects constant. It may also be expressed in a concise form, permitting a better understanding of mechanisms whereas the local-effect approach may not allow a simple expression of the partial equilibrium effect, depending on the structure of matrix $\boldsymbol{\gamma}$. The advantage of the local-effect approach lies in its connection to Theorem 1 in Allen et al. (2024). In the theorem, it is the $|\boldsymbol{\alpha}|$ in the local approach (i.e. equation (7)) that is used to provide both sufficient and "globally" necessary conditions for uniqueness. "Globally" necessary means that when $\rho(|\boldsymbol{\alpha}|) > 1$, then for any matrix of elasticities whose absolute values is equal to $|\boldsymbol{\alpha}|$ there exists geographies where there is multiplicity under the model in (7). Such results suggests that the specification in equation (7) and the associated $\boldsymbol{\alpha}$ may be close to, if not the "best" one, in terms of guaranteeing Neumann representation and reducing the spectral radius, although we do not provide theoretical results in this regard.

2.1 Exact Counterfactual Analysis

Exact counterfactual analysis in the constant-elasticity spatial family all follow the same structure and data requirements, and these data requirements are identical to those of first-order analysis.

Denoting $\hat{\cdot}$ to denote exact changes: $\hat{x} = \frac{x'}{x}$ where x' denotes a counterfactual and x denotes the baseline quantity, the system of equations for exact changes in equilibrium quantities can be written with the same structure as equation (7):

$$\hat{y}_{ih} = \sum_{j \in \mathcal{N}} \left[\mathbf{X}_{(h)}\right]_{ij} \hat{K}_{ijh} \prod_{h' \in \mathcal{H}} \hat{y}_{jh'}^{\alpha_{hh'}} \tag{18}$$

where $\left[\mathbf{X}_{(h)}\right]_{ij}$ is the i, j element of the share matrix $\mathbf{X}_{(h)}$ defined in the previous section.

Given the same structure as before, where K_{ijh} has been replaced by $X_{ijh}\hat{K}_{ijh}$ and other variables now have $\hat{\cdot}$, this system has similar existence and uniqueness properties as the system in (7). Once transformed variables have been computed, then the desired quantities of the model can be computed as $\ln(\mathbf{x}) = \tilde{\mathbf{\Gamma}}^{-1} \ln(\mathbf{y})$, where we note that in the exact representation, we are applying the matrix to the logarithm of variables instead of applying them directly.

We also note that an exact version of "partial equilbrium effect" can be expressed, which may be useful when considering large changes to the geography $\hat{K}_{ijh}$. The exact counterpart to the *local partial equilibrium effect* defined using equation (18), holds fixed the right-hand-side equilibrium variables in other locations: $\hat{y}_{jh'} = 1$. We denote the local partial equilibrium quantity as $\hat{y}^{LP}$.

$$\hat{y}_{ih}^{LP} = \sum_{j \in \mathcal{N}} X_{ijh} \hat{K}_{ijh} \tag{19}$$

and again, the desired variables can be computed as $\ln(\hat{\mathbf{x}}^{LP}) = \tilde{\mathbf{\Gamma}}^{-1} \ln(\hat{\mathbf{y}}^{LP})$

Similarly, the exact *single-variable partial equilibrium effect* can be expressed by holding constant all variables in other locations as well as the other endogenous variables in the current location. Like in the first-order case, the advantage of this approach is that one need not apply a matrix inverse and so a closed-form expression is more easily obtained. Manipulation of equation (3) leads to the following expression, and setting $\hat{x}_{ih'} = 1$ if $h' \neq h$ as well as $\hat{x}_{jh'} = 1$ for all j, h on the right-hand-side leads to the following explicit expression for our desired variable x_{ih}. We denote this variables with the superscript: x^{SP}.

$$\hat{x}_{ih}^{SP} = \left[\sum_{j \in \mathcal{N}} X_{ijh} \hat{K}_{ijh} \right]^{\frac{1}{\gamma_{hh} - \rho_{hh}}} \tag{20}$$

.

The exact representation has the advantage that it does not suffer from first-order approximation issues. Analysis of higher-order effects shows that the first-order effect amounts to an approximation of a log of averages with the average of logs. Hence, the concavity of the logarithm function leads the first-order approximation to underestimate an increase in endogenous variables while overestimating the magnitude of a decrease in endogenous variables. The disadvantage is that we do not present a Neumann series representation and therefore properties of the exact-hat expression (e.g. its size relative to the general equilibrium effects) are not as easily known. We explore these points in our quantitative application.

3 Partial and General Equilibrium Effects for Tokyo's Train System

What is the impact of improvements to commuting infrastructure on where people live and work?

We apply our comparative-statics results to a simple gravity-based commuting model to answer this question. We consider a simplified version of the model in Ahlfeldt et al. (2015), discussed in Allen and Arkolakis (2022). In contrast to Ahlfeldt et al. (2015), this version omits floorspace from the model. As a result, congestion forces arise explicitly from spillovers rather than from the decreasing returns implied by a location-specific fixed resource (land). We show in Section A that the comparative statics are applicable to any commuting model featuring a gravity equation that is log-linear in employment and population. We also show that the model is isomorphic to a model with floorspace. If the amount of land in each location allocated to commercial and residential uses is fixed exogenously, then the reduced-form equations can readily accommodate floorspace in production. If, in addition, preferences for residential housing takes the quasilinear form, then the model can accommodate residential housing. More generally, a wide class of commuting models can be incorporated into the structure of (1), provided land is exogenous and separated between commercial and residential uses.

Indirect utility for an individual $\omega \in [0, 1]$ living in location i and working in location j is given by

$$V_{ij}(\omega) = \frac{U_i w_j \epsilon_{ij}(\omega)}{d_{ij}},$$

where $\epsilon_{ij}(\omega) \sim$ Fréchet(θ), w_j is the wage in workplace j, U_i is the amenity in residence i, and d_{ij} is the commuting cost from i to j. Each location j produces a homogeneous and costlessly traded good with constant-returns-to-scale production and productivity A_j, using only labor: Amenities are a function of an exogenous component $\bar{U}_i$ and residential population: $U_i = \bar{U}_i L_{Ri}^{\beta}$. Productivity is similarly a function of an exogenous component $\bar{A}_j$ and local employment: $A_j = \bar{A}_j L_{Fj}^{\alpha}$. Setting the price of the final good as the numéraire, wages equal productivity: $w_j = A_j = \bar{A}_j L_{Fj}^{\alpha}$.

The Fréchet distribution implies that commuter flows—the number of commuters between residence i and workplace j—are

$$L_{ij} = \bar{L} \frac{\bar{U}_i^{\theta} L_{Ri}^{\beta\theta} \bar{A}_j^{\theta} L_{Fj}^{\alpha\theta}}{\sum_i \sum_j \bar{U}_i^{\theta} L_{Ri}^{\beta\theta} \bar{A}_j^{\theta} L_{Fj}^{\alpha\theta}}, \tag{21}$$

where $\bar{L}$ is total employment in the economy and may or may not be exogenous, depending on population-mobility constraints.

Commuter market clearing requires that residential population in each location equals the sum of outbound commuter flows, and that employment in each location equals the sum of inbound commuter flows:

$$\sum_i L_{ij} = L_{Fj}, \tag{22}$$

$$\sum_j L_{ij} = L_{Ri}. \tag{23}$$

Section A.2 uses equations (21)–(23) to describe equilibrium as a system of $2N$ equations in $2N$ unknowns:

$$\tilde{L}_{Ri} = \sum_j d_{ij}^{-\theta} \bar{U}_i^{\theta} \bar{A}_j^{\theta} \tilde{L}_{Fj}^{\alpha\theta} \tilde{L}_{Ri}^{\beta\theta}, \tag{24}$$

$$\tilde{L}_{Fj} = \sum_i d_{ij}^{-\theta} \bar{U}_i^{\theta} \bar{A}_j^{\theta} \tilde{L}_{Ri}^{\beta\theta} \tilde{L}_{Fj}^{\alpha\theta}, \tag{25}$$

where $\tilde{L}_{Ri} \equiv \kappa_R L_{Ri}$ and $\tilde{L}_{Fj} \equiv \kappa_F L_{Fj}$ are unknowns, and κ_R, κ_F are endogenous scalars that depend on population-mobility constraints.

From Allen et al. (2024), the sufficient condition for uniqueness implies that when congestion forces are strong enough (here, when $\theta\alpha \in (-\frac{1}{2}, \frac{1}{2})$ and $\theta\beta \in (-\frac{1}{2}, \frac{1}{2})$), the system has a unique solution for $\tilde{L}_{Ri}, \tilde{L}_{Fj}$, and hence solutions for L_{Ri}, L_{Fj} that are unique up to scale.

Definition. Given parameters $\{\theta, \alpha, \beta\}$, exogenous characteristics and commuting costs $\{\bar{\mathbf{A}}, \bar{\mathbf{U}}, \mathbf{D}\}$, an equilibrium is defined by the set of endogenous outcomes $\{\mathbf{L}, \mathbf{L_R}, \mathbf{L_F}, \bar{L}\}$, determined up to scale, that satisfy (21)–(23).

An implication of the uniqueness-up-to-scale result is that relative population between two locations (e.g., L_{R1}/L_{R2}) is invariant to the total city population, conditional on primitives $\{\bar{\mathbf{A}}, \bar{\mathbf{U}}, \mathbf{D}\}$. We exploit this property throughout our analysis.

We begin by considering the exact general-equilibrium impact of the train system. Figure 2 shows the changes in population and employment, while Figure 1 shows the corresponding density gradients. We find that Tokyo's train system centralizes both employment and population, with steeper gradients when trains are included. Removing trains from today's equilibrium reduces central Tokyo's share (within 20 km of the Imperial Palace) of population by 10% and of employment by 31%. The employment effect mirrors earlier studies showing that transportation improvements lead to employment concentration in city centers, whereas the centralizing impact on population has few empirical precedents.

Tokyo is one of the densest cities in the world (see Table 1), with roughly 40 million residents—about twice that of the New York–Newark metro area—living in an area about 30% smaller. Understanding why our model implies strong centralization may shed light on both Tokyo's density and the dynamics of other cities with large train networks and dense centers, such as New York City.

Suppose we consider a series of small changes to transportation costs $d \ln d_{ij}$. We are interested in the total differential:

$$d \ln L_{Rl} = \sum_{i,j} \frac{\partial \ln L_{Rl}}{\partial \ln d_{ij}} d \ln d_{ij}.$$

We must first determine $\frac{\partial \ln L_{Rl}}{\partial \ln d_{ij}}$, i.e., the change in population resulting from a change in commuting costs $d \ln d_{ij}$.

We apply the comparative-statics results from Section 2 to describe the impact of a change in commuting costs on the normalized variables $\tilde{L}_{Ri}$ and $\tilde{L}_{Fj}$, expressed as a $2N \times N^2$ matrix:

Theorem 3. *Given a solution $\tilde{\mathbf{L}}_{\mathbf{R}}$ and $\tilde{\mathbf{L}}_{\mathbf{F}}$ to the system of equations* (24)–(25)*, the elasticities of relative population and relative employment with respect to commuting costs are given by the following expression, whenever the matrix $\mathbf{A}$ is full rank:*

$$\begin{bmatrix} \dfrac{\partial \ln \tilde{\mathbf{L}}_{\mathbf{R}}}{\partial \ln \mathbf{d}} \\ \\ \dfrac{\partial \ln \tilde{\mathbf{L}}_{\mathbf{F}}}{\partial \ln \mathbf{d}} \end{bmatrix} = -\mathbf{A}^{-1}\mathbf{T}. \tag{26}$$

where

$$\mathbf{A} \equiv \begin{bmatrix} (1-\beta\theta)\mathbf{I} & -\alpha\theta \dfrac{\mathbf{L}}{\mathbf{L}_{\mathbf{R}}} \\ \\ -\beta\theta \dfrac{\mathbf{L}^{\top}}{\mathbf{L}_{\mathbf{F}}} & (1-\alpha\theta)\mathbf{I} \end{bmatrix}, \qquad \mathbf{T} \equiv \begin{bmatrix} \theta \left(\dfrac{\mathbf{L}}{\mathbf{L}_{\mathbf{R}}} \otimes \mathbf{1} \right) \circ (\mathbf{1} \otimes \mathbf{I}) \\ \\ \theta \left(\mathbf{1} \otimes \dfrac{\mathbf{L}^{\top}}{\mathbf{L}_{\mathbf{F}}} \right) \circ (\mathbf{I} \otimes \mathbf{1}) \end{bmatrix}. \tag{27}$$

Here $\mathbf{1}$ is a row vector of ones, $\mathbf{I}$ is the identity matrix, $\circ$ denotes the Hadamard product, and $\otimes$ denotes the Kronecker product. $\tilde{\mathbf{L}}_{\mathbf{R}} \equiv \kappa_R \mathbf{L}_{\mathbf{R}}$ and $\tilde{\mathbf{L}}_{\mathbf{F}} \equiv \kappa_F \mathbf{L}_{\mathbf{F}}$ are column vectors, where κ_R, κ_F are endogenous scalars that depend on population-mobility constraints. $\mathbf{L}$ is the commuter-flow matrix with entries $L_{ij} = [\mathbf{L}]_{ij}$, $\mathbf{L}_F$ is the employment vector $L_{Fj} = [\mathbf{L}_F]_j$, and $\mathbf{L}_R$ is the residential-population vector $L_{Ri} = [\mathbf{L}_R]_i$. $\mathbf{A}$ is a $2N \times 2N$ matrix and $\mathbf{T}$ is a $2N \times N^2$ matrix.

Specifically:

$$-\frac{\partial \ln \tilde{L}_{Rl}}{\partial \ln d_{ij}} = A^{-1}_{li}\,\theta \frac{L_{ij}}{L_{Ri}} + A^{-1}_{l,N+j}\,\theta \frac{L_{ij}}{L_{Fj}}, \tag{28}$$

$$\tag{29}$$

$$-\frac{\partial \ln \tilde{L}_{Fl}}{\partial \ln d_{ij}} = A^{-1}_{N+l,i}\,\theta \frac{L_{ij}}{L_{Ri}} + A^{-1}_{N+l,N+j}\,\theta \frac{L_{ij}}{L_{Fj}}, \tag{30}$$

where $A^{-1}_{li} \equiv [\mathbf{A}^{-1}]_{li}$ *is the* (l,i) *element of* $\mathbf{A}^{-1}$.

In our notation from Section 2, we have

$$\boldsymbol{\gamma} = \begin{bmatrix} (1-\beta\theta) & 0 \\ & \\ 0 & (1-\alpha\theta) \end{bmatrix}, \qquad \tilde{\boldsymbol{\Gamma}} = \begin{bmatrix} (1-\beta\theta)\mathbf{I} & \mathbf{0} \\ & \\ \mathbf{0} & (1-\alpha\theta)\mathbf{I} \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 & \alpha\theta \\ & \\ \beta\theta & 0 \end{bmatrix},$$

$$\boldsymbol{\alpha} = \begin{bmatrix} 0 & \frac{\alpha\theta}{1-\alpha\theta} \\ & \\ \frac{\beta\theta}{1-\beta\theta} & 0 \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} \frac{\mathbf{L}}{\mathbf{L_R}} & \mathbf{0} \\ & \\ \mathbf{0} & \frac{\mathbf{L}^\top}{\mathbf{L_F}} \end{bmatrix}, \qquad \boldsymbol{\alpha}_X = \begin{bmatrix} 0 & \frac{\alpha\theta}{1-\alpha\theta}\frac{\mathbf{L}}{\mathbf{L_R}} \\ & \\ \frac{\beta\theta}{1-\beta\theta}\frac{\mathbf{L}^\top}{\mathbf{L_F}} & 0 \end{bmatrix}.$$

A number of remarks are in order.

First, writing the comparative statics in terms of normalized variables is useful for studying *relative* changes, since relative changes do not depend on the endogenous scalar: $\frac{\partial \ln(L_{R1}/L_{R2})}{\partial \ln \mathbf{d}} = \frac{\partial \ln(\tilde{L}_{R1}/\tilde{L}_{R2})}{\partial \ln \mathbf{d}}$, and therefore do not depend on population mobility constraints.

How should one interpret the normalized results when we are not comparing two locations? One interpretation is to consider a hypothetical location within the city (i.e., with free mobility) that has positive population and employment but is in commuting autarky (i.e., no commuting in or out). Any comparative static on normalized variables can then be viewed as relative to this hypothetical location.

Second, the comparative statics are written in terms of observables, $\frac{\mathbf{L}}{\mathbf{L_F}}$, $\frac{\mathbf{L}}{\mathbf{L_R}}$, and parameters $\alpha\theta$, $\beta\theta$. The matrix $\frac{\mathbf{L}}{\mathbf{L_R}}$ gives commuting probabilities conditional on residence, while $\frac{\mathbf{L}}{\mathbf{L_F}}$ gives commuting probabilities conditional on workplace. The terms $\alpha\theta$ and $\beta\theta$ are the productivity and residential-congestion parameters scaled by the Fréchet parameter θ. Thus, given parameters and the commuter-flow matrix $\mathbf{L}$, the comparative statics can be readily computed using a matrix inversion. Knowledge of the full model primitives is not required. Since commuter-flow data are commonly available for urban areas, this is a useful result.

Equations (28) and (30) highlight that relative changes in population and employment are each a sum of two terms. Consider the residential-population equation (28). The first term, $A_{li}^{-1}\theta\frac{L_{ij}}{L_{Ri}}$, which we call the *residential effect*, is proportional to the exposure of residents in location i to the commuting improvement from i to j, measured by $\frac{L_{ij}}{L_{Ri}}$, the share of commuters from i who work in j. The second term, $A_{l,N+j}^{-1}\theta\frac{L_{ij}}{L_{Fj}}$, which we call the *firm effect*, is proportional to the exposure of employment in location j to the commuting improvement i to j, measured by $\frac{L_{ij}}{L_{Fj}}$, the share of workers in j who commute from i.

As we will see, the residential effect captures the general-equilibrium impact on location l from residents of i having better access to firms in j, while the firm effect captures the general-equilibrium impact on l from firms in j having better access to residents in i.

To better understand the comparative statics, we implement our Neumann series following Section 2.[12]

Theorem 4. *The matrix inverse of* $\mathbf{A}$ *can be expressed as a Neumann series whenever the spectral radius of* $\boldsymbol{\alpha}_X$ *is less than one:*

$$\mathbf{A}^{-1} = \sum_{k=0}^{\infty} \tilde{\boldsymbol{\Gamma}}^{-1}\boldsymbol{\alpha}_X^k = \sum_{k=0}^{\infty} \underbrace{\begin{bmatrix} \frac{1}{1-\beta\theta}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \frac{1}{1-\alpha\theta}\mathbf{I} \end{bmatrix}}_{\tilde{\boldsymbol{\Gamma}}^{-1}} \underbrace{\begin{bmatrix} 0 & \frac{\alpha\theta}{1-\alpha\theta}\frac{\mathbf{L}}{\mathbf{L_R}} \\ \frac{\beta\theta}{1-\beta\theta}\frac{\mathbf{L^T}}{\mathbf{L_F}} & 0 \end{bmatrix}}_{\boldsymbol{\alpha}_X}{}^{k} \tag{31}$$

Thus, the comparative statics on residential population can be expressed as

$$\frac{\partial \ln \tilde{L}_{Rl}}{\partial \ln d_{ij}} = \frac{-\theta L_{ij}}{L_{Ri}} \underbrace{\left(\tilde{A}_{li}^0 + \tilde{A}_{li}^1 + \cdots\right)}_{A_{li}^{-1}} + \frac{-\theta L_{ij}}{L_{Fj}} \underbrace{\left(\tilde{A}_{l,N+j}^0 + \tilde{A}_{l,N+j}^1 + \cdots\right)}_{A_{l,N+j}^{-1}}, \tag{32}$$

where $\tilde{A}_{li}^k$ *is the* (l, i) *element of* $\tilde{\boldsymbol{\Gamma}}^{-1}\boldsymbol{\alpha}_X^k$.[13]

The terms $\boldsymbol{\alpha}_X^k$ *decay according to*

$$\|\boldsymbol{\alpha}_X^k\|_\infty \leq \zeta^k \quad \text{for all } k = 1, 2, \ldots,$$

where $\zeta \equiv \max\left\{\left|\frac{\alpha\theta}{1-\alpha\theta}\right|, \left|\frac{\beta\theta}{1-\beta\theta}\right|\right\}$ *and* $\|\cdot\|_\infty$ *is the maximum-row-sum norm. The spectral radius is*

$$\rho(\boldsymbol{\alpha}_X) = \gamma \equiv \sqrt{\left|\frac{\alpha\theta\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\right|}.$$

This formulation allows us to decompose the two terms $A_{li}^{-1}\theta\frac{L_{ij}}{L_{Ri}}$ and $A_{l,N+j}^{-1}\theta\frac{L_{ij}}{L_{Fj}}$ (and

[12]See Allen et al. (2020) and Allen and Arkolakis (2025) for a similar Neumann-series expansion in trade and geography models.

[13]Here the superscript k in $\tilde{A}_{li}^k$ indexes propagation degree, not exponentiation.

their sum) into network-propagation effects, where each term isolates the population response to a shock transmitted through the network with different propagation length. It also makes clear that the k-degree impact arises recursively: the k-degree effect reflects transmission of the $(k-1)$-degree response through the network, via equations (24)–(25).

It is useful to define

$$\mathbf{B}_X \equiv \text{bdiag}\Big(\mathbf{X}_{(1)}, \ldots, \mathbf{X}_{(H)}\Big)(\mathbf{B} \otimes \mathbf{I}_N) = \begin{bmatrix} 0 & \alpha\theta \frac{\mathbf{L}}{\mathbf{L_R}} \\ \beta\theta \frac{\mathbf{L^T}}{\mathbf{L_F}} & 0 \end{bmatrix},$$

which implies $\boldsymbol{\alpha}_X = \mathbf{B}_X \tilde{\boldsymbol{\Gamma}}^{-1}$.

The k^{th} *degree residential effect* and k^{th} *degree firm effect* with respect to population l (L_{Rl}) are defined as $-\tilde{A}^k_{li}\theta\frac{L_{ij}}{L_{Ri}}$ and $-\tilde{A}^k_{l,N+j}\theta\frac{L_{ij}}{L_{Fj}}$, respectively, where $\tilde{A}^k_{li}$ is the (l,i) element of $\tilde{\boldsymbol{\Gamma}}^{-1}\boldsymbol{\alpha}^k_X$.

Thus, the k^{th} term of the comparative static is equal to the sum of the k^{th}-degree residential and firm effects:

$$\left[\overline{D}^k_{d_{ij}} \ln \mathbf{x}^*\right]_{Rl} = -\tilde{A}^k_{li}\theta\frac{L_{ij}}{L_{Ri}} \; + \; -\tilde{A}^k_{l,N+j}\theta\frac{L_{ij}}{L_{Fj}}.$$

where $\mathbf{x}^* = [\tilde{L}_{R1}, \ldots, \tilde{L}_{RN}, \tilde{L}_{F1}, ..., \tilde{L}_{FN}]'$ and $\left[\overline{D}^k_{d_{ij}} \ln \mathbf{x}^*\right]_{Rl}$ denotes the l^{th} element of the "R" block (i.e. the first block) of the vector $\overline{D}^k_{d_{ij}} \ln \mathbf{x}^*$.

Consider for example a setting where commuting costs fall between two locations, 1 and 2, by $-d\ln d_{12}$, improving the commute from location 1 to 2.

The first term of the residential-effect series (the zeroth-degree effect) is $\tilde{A}^0_{11}\theta\frac{L_{12}}{L_{R1}} = [\tilde{\boldsymbol{\Gamma}}^{-1}]_{11}\theta\frac{L_{12}}{L_{R1}} = \frac{1}{1-\beta\theta}\theta\frac{L_{12}}{L_{R1}}$. It captures the impact of the commuting improvement on locations directly connected to the improvement (i.e., zero degrees of separation from the shock), prior to interaction with other locations through the commuting network. It consists of two parts: the direct residential propagation effect and the local equilibrium response. The *direct propagation effect* is the immediate impact on the equilibrium variable (in this case, population in location 1) that results from the propagation shock (zeroth degree) while holding the congestion force in that location constant. In our example, the zeroth-degree direct residential effect is $\theta\frac{L_{12}}{L_{R1}}$. The local equilibrium response is the adjustment of the equilibrium variable in response to the change in congestion in that location (e.g., in location 1), which in this case is captured by the factor $\frac{1}{1-\beta\theta}$ in the diagonal matrix $\tilde{\boldsymbol{\Gamma}}^{-1}$. The total zeroth-degree effect is then $\tilde{A}^0_{11}\theta\frac{L_{12}}{L_{R1}} = \frac{1}{1-\beta\theta}\theta\frac{L_{12}}{L_{R1}}$.

On the other hand, the zeroth-degree firm effect is zero because the direct effect on firms in location 2 only affects those firms (hence the off-diagonal elements of $\tilde{\boldsymbol{\Gamma}}^{-1}$ are zero) and requires an additional degree of separation to reach residential location 1. Thus, the zeroth degree propation effect is equal to $\frac{1}{1-\beta\theta}\frac{L_{12}}{L_{R1}}$.

Consider next the first-degree effects: $\tilde{A}^1_{11}\theta\frac{L_{12}}{L_{R1}} = [\tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}_X]_{11}\theta\frac{L_{12}}{L_{R1}} = [\tilde{\mathbf{\Gamma}}^{-1}\mathbf{B}_X\tilde{\mathbf{\Gamma}}^{-1}]_{11}\theta\frac{L_{12}}{L_{R1}}$, and $\tilde{A}^1_{1,N+2}\theta\frac{L_{12}}{L_{F2}} = [\tilde{\mathbf{\Gamma}}^{-1}\mathbf{B}_X\tilde{\mathbf{\Gamma}}^{-1}]_{1,N+2}\theta\frac{L_{12}}{L_{R1}}$. Following the zeroth-degree shock to location-1 residents, the first-degree residential effect propagates to all firm locations, each one experiencing a direct shock that scales the zeroth-degree shock by $\alpha\theta\frac{L_{1j}}{L_{Fj}}$ for each workplace j, represented by matrix $\mathbf{B}_X$. However, residential locations do not experience such a shock because residential spillovers are local and do not extend to other residential locations (hence the diagonal blocks of $\mathbf{B}_X\tilde{\mathbf{\Gamma}}^{-1}$ are zero). Therefore, the first-degree residential effect is zero.

In contrast, the first-degree firm effect is nonzero. Following the zeroth-degree shock to location-2 employment, the first-degree firm effect propagates to all residential locations, each one experiencing a direct shock that scales the zeroth-degree shock ($\tilde{A}^0_{1,N+2}\theta\frac{L_{12}}{L_{F2}} = \frac{1}{1-\alpha\theta}\theta\frac{L_{12}}{L_{F2}}$) by the factor $\beta\theta\frac{L_{i2}}{L_{Ri}}$ for each residential location i (represented by the lower-left block of $\mathbf{B}_X$). The local equilibrium response in residential location 1 then scales this shock by $\frac{1}{1-\beta\theta}$ (the first block of $\tilde{\mathbf{\Gamma}}^{-1}$), producing a total first-degree effect of $\tilde{A}^1_{1,N+2}\theta\frac{L_{12}}{L_{F2}} = [\tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}_X]_{1,N+2}\theta\frac{L_{12}}{L_{F2}} = \frac{\alpha\theta}{(1-\beta\theta)(1-\alpha\theta)}\theta\frac{L_{12}}{L_{F2}}$.

By now, the recursive nature of the k^{th}-degree is clear. Given the $(k-1)$-degree residential effects across all locations, $\tilde{A}^{k-1}_{ij}\theta\frac{L_{12}}{L_{R1}} = [\tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}^{k-1}_X]_{ij}\theta\frac{L_{12}}{L_{R1}}$, the k^{th}-degree direct residential effect is a linear combination of the $(k-1)$-degree effects weighted by $\mathbf{B}_X$, leading to a direct impact of $\mathbf{B}_X[\tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}^{k-1}_X]_{ij}\theta\frac{L_{12}}{L_{R1}}$. The local equilibrium response then scales this by $\tilde{\mathbf{\Gamma}}^{-1}$, leading to a total k^{th}-degree residential effect of $[\tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}_X\boldsymbol{\alpha}^{k-1}_X]_{11}\theta\frac{L_{12}}{L_{R1}} = [\tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}^{k}_X]_{11}\theta\frac{L_{12}}{L_{R1}}$. Analogous expressions hold for the firm effect, and also for comparative statics on employment (equation (30)).

In summary, $\mathbf{B}_X$ mediates the direct propagation effect between degrees of separation, while $\tilde{\mathbf{\Gamma}}^{-1}$ captures the local equilibrium response. The k^{th}-degree term of the comparative static is thus the residual impact of the initial shock after dissipation through the network (via equations 24 and 25) k times. The total impact is the sum of all k-degree comparative static terms. The following theorem summarizes these results and provides bounds on the relative importance of the partial equilibrium effect relative to the full general equilibrium effect, for a location whose residents receive improved commuting access to a workplace.

Definition. The *partial equilibrium impact on population* for location l for a change in commuting costs d_{ij} is the partial elasticity of normalized population, $\tilde{L}_{Rl}$ with respect to d_{ij}, holding constant normalized employment in all locations (including l): $\left.\frac{\partial \ln \tilde{L}_{Rl}}{\partial \ln d_{ij}}\right|_{\tilde{L}_{Fj}\text{ fixed for all } j}$.[14]

[14]In this application, the right-hand side variable (employment) is indeed fixed in all locations, rather

We note that in this application, the single-variable partial equilibrium effect and the local partial equilibrium effect correspond to the same expression, because each equation (24) and (25) contains only one "local" variable and hence $\boldsymbol{\gamma}$ is a diagonal matrix. Thus the log of the composite variable is proportional to the log of the endogenous variable of interest. We therefore do not need to consider alternative changes of variables to obtain simple expressions for the partial equilibrium effect, as we might in a more complex model (where $\boldsymbol{\gamma}$ is not diagonal).

Theorem 5. ***Partial-equilibrium Impact on Population***
(i.) The partial equilibrium effect is the zeroth-degree term of the comparative static and is equal to $\frac{\partial \ln \tilde{L}_{Rl}}{\partial \ln d_{ij}}\Big|_{\tilde{L}_{Fj} \text{ fixed for all } j} = \frac{-\theta L_{ij}}{L_{Ri}} \frac{1}{1-\beta\theta}$ *if* $l = i$ *and 0 otherwise.*
(ii.) Applying the bounds in Theorem 1 (iv) implies that for any small change $d \ln d_{ij}$*, the magnitude of the remainder term is less than* $\max\left\{\left|\frac{1}{1-\beta\theta}\right|, \left|\frac{1}{1-\alpha\theta}\right|\right\} \max\left\{\frac{\theta L_{ij}}{L_{Ri}}, \frac{\theta L_{ij}}{L_{Fj}}\right\} \frac{\|\boldsymbol{\alpha}_X\|}{1-\|\boldsymbol{\alpha}_X\|}$ *where* $\|\boldsymbol{\alpha}_X\| = \max\left\{\left|\frac{\alpha\theta}{1-\alpha\theta}\right|, \left|\frac{\beta\theta}{1-\beta\theta}\right|\right\}$.
(iii.) If $l = i$ *and* $\alpha\beta > 0$*, the partial equilibrium effect is at least* $1 - \gamma^2$ *of the residential effect* $A_{li}^{-1}\theta \frac{L_{ij}}{L_{Ri}}$*, and is at least* $\frac{1-\gamma^2}{1+\frac{L_{iR}}{L_{jF}}\left|\frac{\alpha\theta}{1-\alpha\theta}\right|}$ *of the full general equilibrium effect* $\frac{\partial \ln \tilde{L}_{Rl}}{\partial \ln d_{ij}}$*, where* $\gamma = \sqrt{\left|\frac{\alpha\theta\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\right|}$.

Proof. See Section C. □

Section C shows that the block off-diagonal structure of matrix $\boldsymbol{\alpha}_X$ in this problem allows us to provide bounds beyond those in Section 2. When the elasticities α and β have different signs, the bounds improve further due to the alternating impacts of each shock.[15]

In Allen and Arkolakis (2022), the authors parametrize $\alpha = -0.12, \beta = -0.1, \theta = 6.83$, based on estimates from Ahlfeldt et al. (2015), such that our conditions on the signs of α, β are satisfied, yielding $\|\boldsymbol{\alpha}_X\| \approx 0.45$, $\gamma \approx 0.43$, and $\gamma^2 \approx 0.18$. With these assumptions, part (ii) implies that whenever one considers a change in commute costs $d \ln d_{ij}$ with $L_{Ri} < L_{Fj}$, the partial equilibrium effect on population L_{Ri} is at least 1.2 times the remainder term and at least 54% of the general equilibrium effect. If we use part (iii), these parameters imply that the partial equilibrium is at least 81% of the residential effect, and if we consider a comparative static where the population in the location in question is equal to employment in the commuting location (i.e., $L_{iR} = L_{jF}$), then the partial equilibrium effect is at least 56% of the full general equilibrium effect.

than the more artificial definition we discussed in Section 3. The reason is that the left-hand side variable (population) does not appear on the right-hand side, so there is no ambiguity about what is being held constant.

[15] Many gravity-related models may share this structure (and therefore similar bounds), because the sender and receiver are the two main interactions and form a closed loop after two propagations.

A limitation of the above theorem is that the comparative-static expression speaks to the impact on "normalized" population, which is less useful in isolation. However, comparisons of this expression across locations are more directly applicable, as we now discuss.

Most transportation improvements increase commuting access in both directions. If commuting infrastructure improves between two locations while commuting costs are unchanged between all other pairs, what is the partial equilibrium impact on the relative populations of these two locations, and how does it compare with the general equilibrium impact?

In other words, if there is a symmetric improvement in commuting between locations 1 and 2, we are interested in the comparative static $-\left[\frac{\partial \ln(L_{R1}/L_{R2})}{\partial \ln d_{12}} + \frac{\partial \ln(L_{R1}/L_{R2})}{\partial \ln d_{21}}\right]$. The following theorem addresses this case when $\alpha, \beta < 0$, which is the assumption in our quantitative analysis.

Definition. Given a series of changes to parameters $\Delta \ln \boldsymbol{\theta} \in \mathbb{R}^M$ where $[\Delta \ln \boldsymbol{\theta}]_m = \Delta \ln \theta_m$, the k^{th} *degree comparative static* for endogenous variable x_{lh} is equal to

$$\overline{\Delta}^k \ln x_{lh} \equiv \mathbf{e}_{lh}' \overline{D}_{\boldsymbol{\theta}}^k \ln \mathbf{x}^* \Delta \ln \boldsymbol{\theta}.$$

where $\mathbf{e}_{lh} \in \mathbb{R}^{NH}$ is the vector with 1 in the l^{th} element of the h^{th} and 0 in all others.

Definition. The k^{th} *degree comparative static* on residential location l for a series of changes to commuting costs $\{d \ln d_{ij}\} \in \mathbb{R}^{N \times N}$ is the sum of the k^{th}-degree individual propagation effects for each change in commuting cost, weighted by $d \ln d_{ij}$: $\sum_{ij} \left(\tilde{A}_{li}^k \theta \frac{L_{ij}}{L_{Ri}} + \tilde{A}_{l,N+j}^k \theta \frac{L_{ij}}{L_{Fj}} \right) d \ln d_{ij}$.

Theorem 6. ***Partial-equilibrium effect for symmetric improvement:*** $-\left[\frac{\partial \ln(L_{Ri}/L_{Rj})}{\partial \ln d_{ij}} + \frac{\partial \ln(L_{Ri}/L_{Rj})}{\partial \ln d_{ji}}\right.$

(i.) For a symmetric transportation improvement between two locations i and j, such that $d \ln d_{ij} = d \ln d_{ji}$, the partial equilibrium effect on the relative populations, $\left.\frac{\partial \ln(L_{Ri}/L_{Rj})}{\partial \ln d_{ij}} + \frac{\partial \ln(L_{Ri}/L_{Rj})}{\partial \ln d_{ji}}\right|_{\tilde{L}_{Fl} \text{ fixed for all } l}$, is equal to $\frac{-\theta}{1-\beta\theta}\left(\frac{L_{ij}}{L_i^R} - \frac{L_{ji}}{L_j^R}\right)$ and corresponds to the difference in zeroth-degree terms of the comparative statics between the two locations.

(ii.) Based on Theorem 1, the magnitude of the general equilibrium remainder is bounded by $2 \times \max\left\{\left|\frac{1}{1-\beta\theta}\right|, \left|\frac{1}{1-\alpha\theta}\right|\right\} \max\left\{\frac{\theta L_{ij}}{L_{Ri}}, \frac{\theta L_{ji}}{L_{Rj}}, \frac{\theta L_{ij}}{L_{Fj}}, \frac{\theta L_{ji}}{L_{Fi}}\right\} \frac{\|\boldsymbol{\alpha}_X\|}{1-\|\boldsymbol{\alpha}_X\|}$ where $\|\boldsymbol{\alpha}_X\| = \max\left\{\left|\frac{\alpha\theta}{1-\alpha\theta}\right|, \left|\frac{\beta\theta}{1-\beta\theta}\right|\right\}$.

(iii.) Given statistics $\frac{L_{ij}}{L_{Ri}}, \frac{L_{ji}}{L_{Rj}}$ as well as workplace conditional probabilities $\frac{L_{ij}}{L_{Fj}}, \frac{L_{ji}}{L_{Fi}}$, and parameters $\alpha\theta, \beta\theta$, a sufficient condition for a population shift from j to i (i.e. $-\left[\frac{\partial \ln(L_{Ri}/L_{Rj})}{\partial \ln d_{ij}} + \frac{\partial \ln(L_{Ri}/L_{Rj})}{\partial \ln d_{ji}}\right] > 0$) is if $\frac{L_{ij}}{L_{Ri}} > \frac{1}{1-2\gamma^2}\left[\frac{L_{ji}}{L_{Rj}} + \left(\frac{L_{ij}}{L_{Fj}} + \frac{L_{ji}}{L_{Fi}}\right)\bar{\alpha}\right]$, whenever $\alpha\theta \in (-1, 0]$, $\beta\theta \in (-1, 0]$, where $\bar{\alpha} \equiv \left|\frac{\alpha\theta}{1-\alpha\theta}\right|$. Equivalently, given statistics L_{ij}, L_{ji} as well as population and employment in both locations $L_{Rj}, L_{Ri}, L_{Fj}, L_{Fi}$, and parameters $\alpha\theta, \beta\theta$, a sufficient condition for a population shift from j to i is if $\frac{L_{ij}}{L_{ji}} > \frac{L_{Ri} L_{Fj}(L_{Fi} + L_{Rj}\bar{\alpha})}{L_{Rj} L_{Fi}(L_{Fj}(1-2\gamma^2) - L_{Ri}\bar{\alpha})}$.

Proof. See Section C.3. □

To put some numbers to the bounds in part (ii.), suppose that $\frac{\theta L_{ij}}{L_{Ri}} = \max\left\{\frac{\theta L_{ij}}{L_{Ri}}, \frac{\theta L_{ji}}{L_{Rj}}, \frac{\theta L_{ij}}{L_{Fj}}, \frac{\theta L_{ji}}{L_{Fi}}\right\}$. Then, with $\alpha = -0.12$, $\beta = -0.1$, the remainder term is less than 1.64 times $\left|\frac{1}{1-\beta\theta}\right| \frac{\theta L_{ij}}{L_{Ri}}$. This information alone is not enough to guarantee whether there will be a population shift, because the remainder term could still be larger in magnitude than the partial-equilibrium effect.

Consider next the bounds in part (iii.), which use the structure of the $\boldsymbol{\alpha}_X$ matrix to produce potentially tighter bounds. Suppose workplace conditional probabilities between two locations equal the residential conditional probability of one location—for example, if $\frac{L_{ij}}{L_{Fj}} = \frac{L_{ji}}{L_{Fi}} = \frac{L_{ji}}{L_{Rj}}$—then we have a population shift from location j to i if $\frac{\frac{L_{ij}}{L_{Ri}}}{\frac{L_{ji}}{L_{Fi}}} > \frac{1}{1-2\gamma}\left[1 + 2\bar{\alpha}\right] \approx 3$, using the same parameters. Under these assumptions, if the share of location 1's residents commuting to location 2 is at least three times the share of location 2 residents commuting to location 1, then a symmetric transportation improvement will shift population from location 2 to location 1.

The theorem establishes that the statistic $\frac{-\theta}{1-\beta\theta}\left(\frac{L_{ij}}{L_i^R} - \frac{L_{ji}}{L_j^R}\right)$ is the leading term of a decaying series, and it offers a way to guarantee the direction of a population shift using readily available statistics.

One view of this sufficient condition is as a proof of concept showing how the matrix structure can yield sharp theoretical predictions. One might think the partial-equilibrium effect should have more predictive power than this result suggests, since many general-equilibrium effects cancel out. In practice, however, we find that the sufficient condition is informative in our application.

With our Tokyo data, we consider all combinations of symmetric improvements $\left[\frac{\partial \ln(L_{Ri}/L_{Rj})}{\partial \ln d_{ij}} + \frac{\partial \ln(L_{Ri}/L_{Rj})}{\partial \ln d_{ji}}\right]$ between pairs of locations $i \neq j$.[16] For these pairs of locations we compute our bounds and find that in 44% of improvements, our sufficient conditions are satisfied and we can guarantee a population increase in one direction or the other.

Next, we run a linear regression of the general-equilibrium effect on our partial-equilibrium statistic, $\frac{\theta}{1-\beta\theta}\left(\frac{L_{ij}}{L_i^R} - \frac{L_{ji}}{L_j^R}\right)$. This regression yields an R^2 of 0.997 and a coefficient of 0.978, close to unity. The scatter plot is shown in Figure 3.

The strong predictive power arises from two factors. First, spatial spillovers are relatively mild, leading to a spectral radius of 0.43, which—as noted earlier—means the partial-equilibrium effect accounts for a large share of the full effect. Second, when we consider

[16] This is an artificial exercise, because any time there is an improvement between two locations there are likely also improvements in connections between other pairs, due to the network structure of transportation and as emphasized by Allen and Arkolakis (2022). Yet we view the exercise as meaningful for isolating the theoretical effect of an improvement between just two locations.

an isolated improvement between two locations, we only need to consider eight general-equilibrium spillover channels. Each direction of improvement (d_{ij} and d_{ji}) has spillovers through both residential and firm effects for each location. The number of spillover channels grows with the square of the number of locations affected, which reduces the predictive power of partial equilibrium when many commuting links change simultaneously.

What does this statistic tell us about Tokyo and the centralizing versus decentralizing mechanisms in commuting models more broadly?

Our first observation is that the sufficient statistic is proportional to the difference in a readily available statistic for two locations: the fraction of each location's residents who commute to the other location. In partial equilibrium, the difference in these two fractions is proportional to the partial-equilibrium effect of a marginal symmetric transport improvement between the two locations on their relative populations. In other words, if a higher proportion of residents commute from one location to another than vice versa, then the location with the higher proportion gains population relative to the other, according to the partial-equilibrium effect.

The implications of this observation are twofold. First, radial improvements in commuting infrastructure (i.e., those radiating from the center to the periphery) tend to decentralize population away from the center whenever commuting from suburb to center is more common than "reverse" commuting (center to suburb), as is typical in many cities.

In the Tokyo context, the average municipality in the periphery, which we define as the region more than 40 km from the Imperial Palace, has a commuting share to the center (within 20 km of the Imperial Palace) that is 20 times larger than that of the reverse flow. Thus, radial improvements in Tokyo tend to shift population to the periphery when considered in isolation. Our bounds guarantee (ignoring general-equilibrium spillovers) that symmetric improvements between peripheral and central locations increase peripheral population in 64% of pairs, while central population increases in only about 1% of pairs. Once we account for general-equilibrium effects of these isolated symmetric improvements, we find that in 91% of center–periphery pairs a symmetric improvement leads to a population shift to the periphery.

A caveat is that transportation infrastructure forms a network: changing one link alters commuting costs across many pairs in non-uniform ways. We refer the reader to Allen and Arkolakis (2022), which provides an elegant methodology for analyzing this problem.

Second, subways and other center-focused transportation improvements may improve access for all commuters but disproportionately benefit residents of the center, leading to population centralization. Consider, for example, a symmetric four-location model arranged in a line: 1–2–3–4, where locations 2 and 3 represent the core and locations 1 and 4 the periphery. Suppose there is no commuting across the entire city (so locations 1 and 4 cannot reach each other). Now suppose commuting infrastructure improves along the core–core link (2–3) as well as along the core–periphery links (1–2 and 3–4). Under common parameterizations of commuting costs, the percentage drop in costs between the cores is at least as large as the drop between core and periphery: $|d \ln d_{23}| \geq |d \ln d_{13}|$.[17] In such an example, the improvement tends to shift population from periphery to center if the share of core-to-core commuting exceeds the share of peripheral-to-distant-core commuting (from 1 to 3 or from 4 to 2).

In the Tokyo context, compared to the average peripheral municipality, the average central municipality is 26 times more likely to commute a location in central Tokyo, and the average suburban municipality (defined as 20-40 km from the Imperial Palace) is 5 times more likely to commute to central Tokyo. Our theory therefore predicts that central-focused transportation improvements will centralize population. Indeed, if we consider a 1% symmetric improvement between central locations (holding within-location costs constant) and a 1% symmetric improvement between central and non-central (peripheral and suburban) locations, the partial-equilibrium effect implies, on average, a 0.6 percentage point population shift toward central Tokyo relative to the periphery and a 0.3 point shift relative to the suburbs. Once general-equilibrium effects are included, which act as a congestion force, the centralizing effect is reduced to 0.3 and 0.2 percentage points, respectively.[18]

In summary, the question of whether or not a transportation system centralizes or decentralizes population can be understood through the lens of these two types of improvements, radial versus central-focused improvements.

Next, we consider the impact on population for a series of changes to commuting costs, namely the impact of Tokyo's entire train system. The first-order impact on normalized population can be expressed as

$$\Delta \tilde{L}_{Rl} = \sum_{i,j} \frac{\partial \ln \tilde{L}_{Rl}}{\partial \ln d_{ij}} \Delta \ln d_{ij},$$

while for the zeroth degree (i.e., partial-equilibrium effects), we can exclude all $i \neq l$ and

[17]For example, if commuting costs are multiplicative across links, $d_{13} = d_{12}d_{23}$, then $d \ln d_{13} = d \ln d_{23}$. If they are a power function of total commute time, $d(t) = t^{\xi}$, then $|d \ln d_{13}| = \left|\frac{\xi t_{23}}{t_{12}+t_{23}} d \ln t_{23}\right| < |\xi d \ln t_{23}| = |d \ln d_{23}|$. More generally, if $d = f(t)$ where t is commute time, this condition is violated only if the elasticity of f grows faster than linearly in t, i.e. $\frac{d}{d \ln t} \frac{d \ln f}{d \ln t} > 1$. An example is $f(t) = e^{t^2}$, a form rarely used and not supported by Tokyo's commuter-flow data.

[18]Even if we include a 1% improvement between all suburban municipalities and peripheral ones, there remains a 0.5% shift in population from the periphery to the center.

write

$$\Delta^P \ln \tilde{L}_{Rl} = \frac{\theta}{1-\beta\theta} \sum_j \frac{L_{lj}}{L_l^R} \Delta \ln d_{lj},$$

where we denote the partial-equilibrium change by Δ^P. Thus, the partial-equilibrium effect on location l is proportional to the weighted average of log changes in commuting costs between location l and other locations, where the weights are the commuting shares to each workplace location.

The exact change in population, $\hat{\tilde{L}}_{Ri}$, is the solution (together with $\hat{L}_{Fj}$) of the following exact-hat system:

$$\hat{\tilde{L}}_{Ri} = \sum_j \frac{L_{ij}}{L_{Ri}} \hat{d}_{ij}^{-\theta} \hat{\tilde{L}}_{Fj}^{\alpha\theta} \hat{\tilde{L}}_{Ri}^{\beta\theta}, \qquad \hat{\tilde{L}}_{Fj} = \sum_i \frac{L_{ij}}{L_{Fj}} \hat{d}_{ij}^{-\theta} \hat{\tilde{L}}_{Ri}^{\beta\theta} \hat{\tilde{L}}_{Fj}^{\alpha\theta}.$$

The *exact* partial-equilibrium effect, where normalized employment is held constant in all locations and which we denote $\hat{L}_{Ri}^P$, can be written as

$$\hat{\tilde{L}}_{Ri}^P = \left[\sum_j \frac{L_{ij}}{L_{Ri}} \hat{d}_{ij}^{-\theta} \right]^{\frac{1}{1-\beta\theta}}. \tag{33}$$

The exact-hat expression has the advantage of giving the exact partial-equilibrium effect rather than just the first-order effect. While we do not characterize its magnitude relative to the remaining general-equilibrium forces, we would expect it to have strong predictive ability, especially when studying the full exact general-equilibrium effect (not only the first order).

We assess the extent to which the first-order ($\Delta \ln \tilde{L}_{Rl} = \sum_{i,j} \frac{\partial \ln L_{Rl}}{\partial \ln d_{ij}} \Delta \ln d_{ij}$) and exact-hat ($\ln \hat{\tilde{L}}_{Ri}^P = \ln \left[\sum_j \frac{L_{ij}}{L_{Ri}} \hat{d}_{ij}^{-\theta} \right]^{\frac{1}{1-\beta\theta}}$) partial-equilibrium effects explain the full general-equilibrium effects. Specifically, we compare the exact and first-order general-equilibrium effects of adding Tokyo's entire train system to the first-order partial-equilibrium effect, which we denote $\Delta \ln L_{Rl}^P = \sum_{i,j} \frac{\partial \ln L_{Rl}}{\partial \ln d_{ij}} \Big|_{\tilde{L}_{Fl} \text{ fixed for all } l} \Delta \ln d_{ij}$.

Because general-equilibrium effects on a particular location is a sum of effects from all improvements including those that are not in direct contact with the location ($l \neq i$), the predictive power of the partial-equilibrium statistic should diminish. Given Tokyo's extensive transport network, we expect these general-equilibrium forces to be large, so we view this exercise as a "lower bound" on the predictive value of partial equilibrium.

Regression of the first-order general-equilibrium effect on the first-order partial-equilibrium effect shows that the partial-equilibrium effect explains 82% of the variation in general-equilibrium first-order population changes due to Tokyo's train system. The

OLS slope coefficient is 0.67, indicating that the partial-equilibrium expression is biased downward because of congestion forces embodied in the two spillover elasticities, both negative.

Regression of the exact general-equilibrium effect on the first-order partial-equilibrium effect does not perform as well, with an R^2 of 0.58. The plot shows how higher-order effects matter: the 45-degree line lies mostly below the logarithmic-like curve, consistent with Jensen's inequality playing a role. Finally, regression of the exact general-equilibrium effect on the exact partial-equilibrium effect yields an R^2 of 0.91 and an OLS coefficient of 0.6. These scatter plots are shown in Figures 4 and 5.

These results indicate that partial-equilibrium measures have strong predictive power (under our parameter choices) even when considering large changes to infrastructure, but their numerical values should not be used at face value for comparative-static analysis.

Figure 2 shows the changes in population and employment in exact counterfactuals of adding and removing the train system, and shows that Tokyo's train system exerts a large centralizing force on both population and employment. In partial equilibrium, central Tokyo experiences a 90% higher growth in residential access than the periphery (calculated using the term in square brackets in the exact-hat partial equilibirium expression, equation 33), translating to an approximately 45% higher change in central Tokyo's population share relative to the periphery. However, 88% of central Tokyo's growth in access was from improved access to other central regions while 12% was from improved access to suburban and periphal regions. These findings, together with our theory suggest that a large part of Tokyo's centralization is due to its dense subway system.

In our final quantification (still to do), we show quantitatively that population centralization in Tokyo is indeed largely due to the dense core subway network. In counterfactuals where we remove subway segments one by one in reverse order of construction, we find each segment contributed significantly to centralization. In contrast, when we do the same exercise for non-subway lines, which are more often commuter lines radiating to the suburbs, we find these lines tended to shift population outward.

4 Conclusion

What is the impact of commuting infrastructure improvements on where people live and work in a city? In the case of Tokyo, we find that the train system has a centralizing impact on both population and employment. Why?

To answer this question, we develop a comparative-statics framework for a wide class of spatial models, of which our commuting model is one example. In addition to generalizing results on comparative statics from the international trade literature to constant-elasticity spatial models, we show that partial-equilibrium analysis provides a useful lens for understanding the mechanisms of such general-equilibrium models.

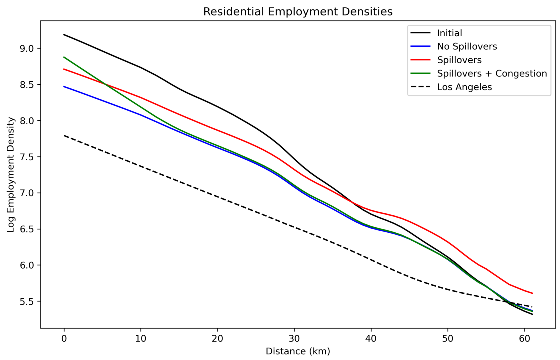
Applying this framework to a canonical urban commuting model, we argue that radial transportation improvements—links connecting the city center to the periphery—tend to shift population outward, while central-focused improvements—links within the city core—tend to shift population inward, even when peripheral locations also benefit from the improved connection. For Tokyo, we find strong evidence (based on sufficient-statistics) that these mechanisms are at work. We further conduct counterfactuals removing subway segments and find, consistent with theory, that the subway has a centralizing effect. Similarly, when we remove non-subway commuter lines, we find, again consistent with theory, that these lines have a decentralizing effect.

In summary, Tokyo's train system centralized both population and employment, driven by its dense subway network concentrated in the city's core, and our theoretical framework explains the mechanisms underlying this outcome.

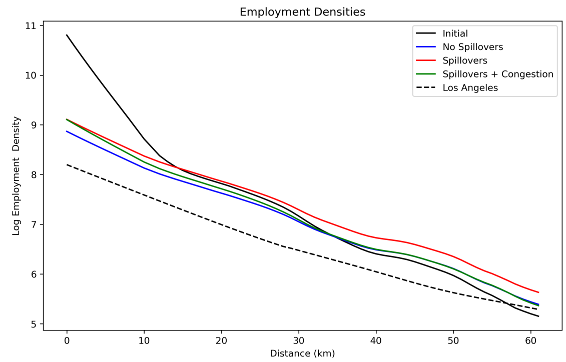
5 Figures

	Tokyo Metro	New York Metro
Population	∼39m	∼20m
GDP	∼$2t USD	∼$2t USD
Area	14,000 sq km	21,000 sq km
Density (people/sq km)	2,600	900
Urban Population	10m	9m (NYC)
Urban Density (people/sq km)	15,000	11,000
Share of Car Commuters	23%	57%
Share of Train Commuters	54%	19%
Average Commute Time	40min	36min
Track Length	∼5,200 km	∼3,500 km[19]

Table 1: Statistics on Two Cities

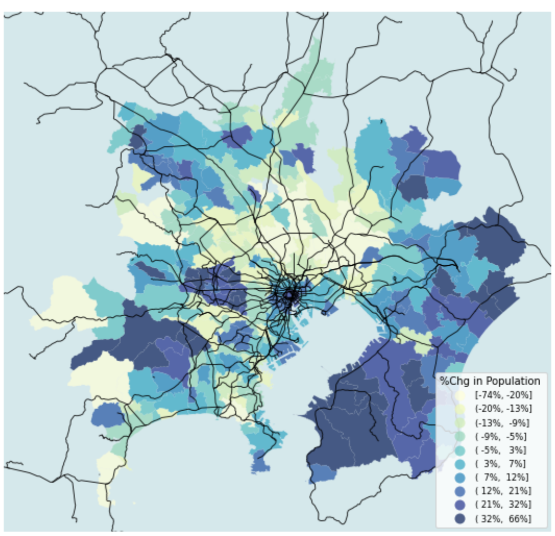


Population Gradient

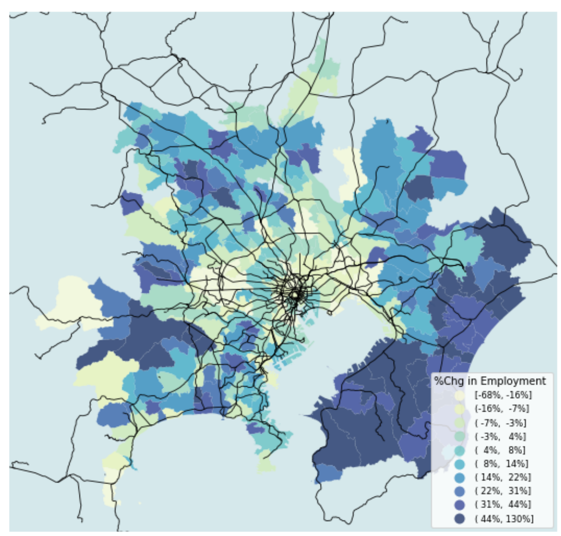


Employment Gradient

Figure 1: (need to edit) Log Population and Employment per Square Kilometer Versus Distance to the Imperial Palace.



Population Gradient



Employment Gradient

Figure 2: (need to edit) Change in Population and Employment from **Removal** of Train System.

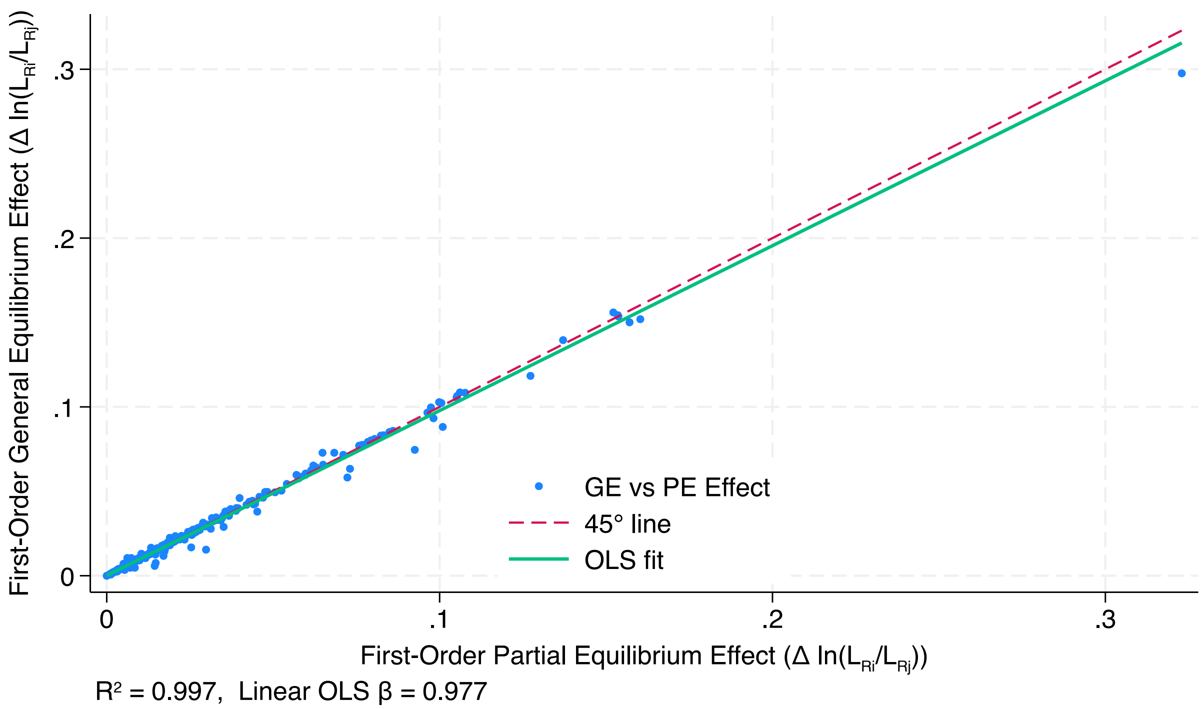


Figure 3: First-order General Equilibrium Effects Versus Partial Equilibrium Effects of an Isolated Symmetric Improvement Between Pairs of Municipalities in Tokyo

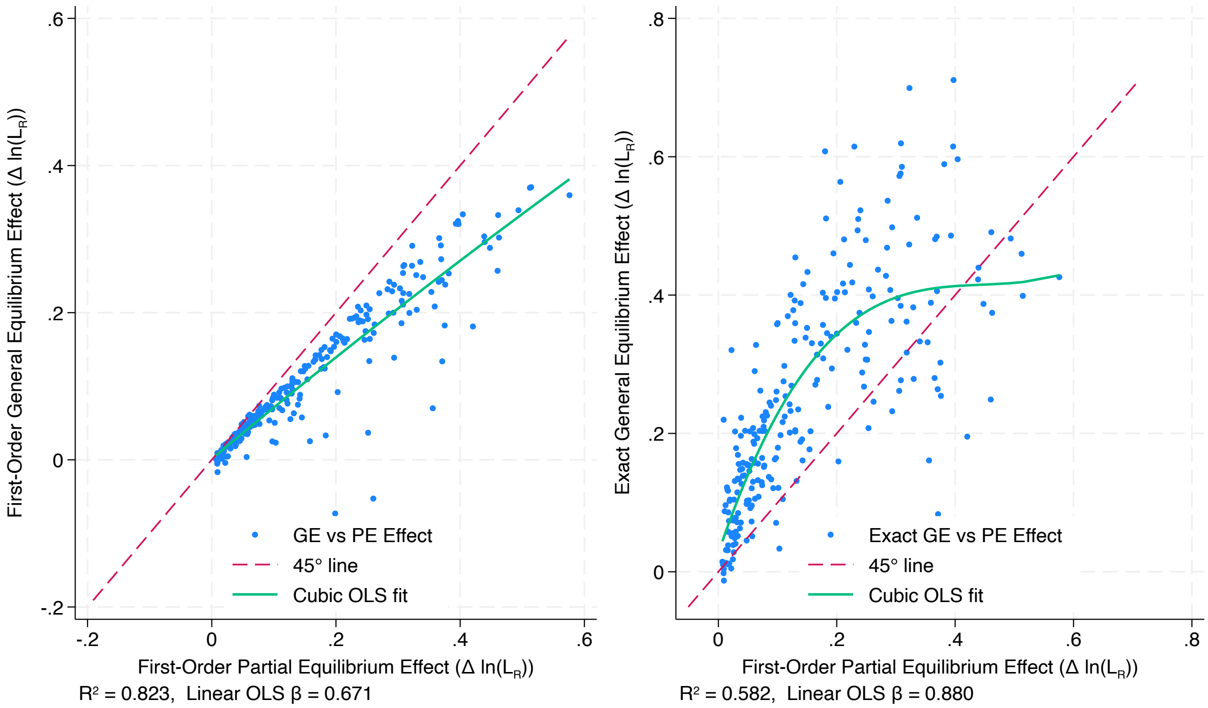


Figure 4: General Equilibrium Effects Versus First-order Partial Equilibrium Effects of Adding Tokyo's Train System Across Municipalities

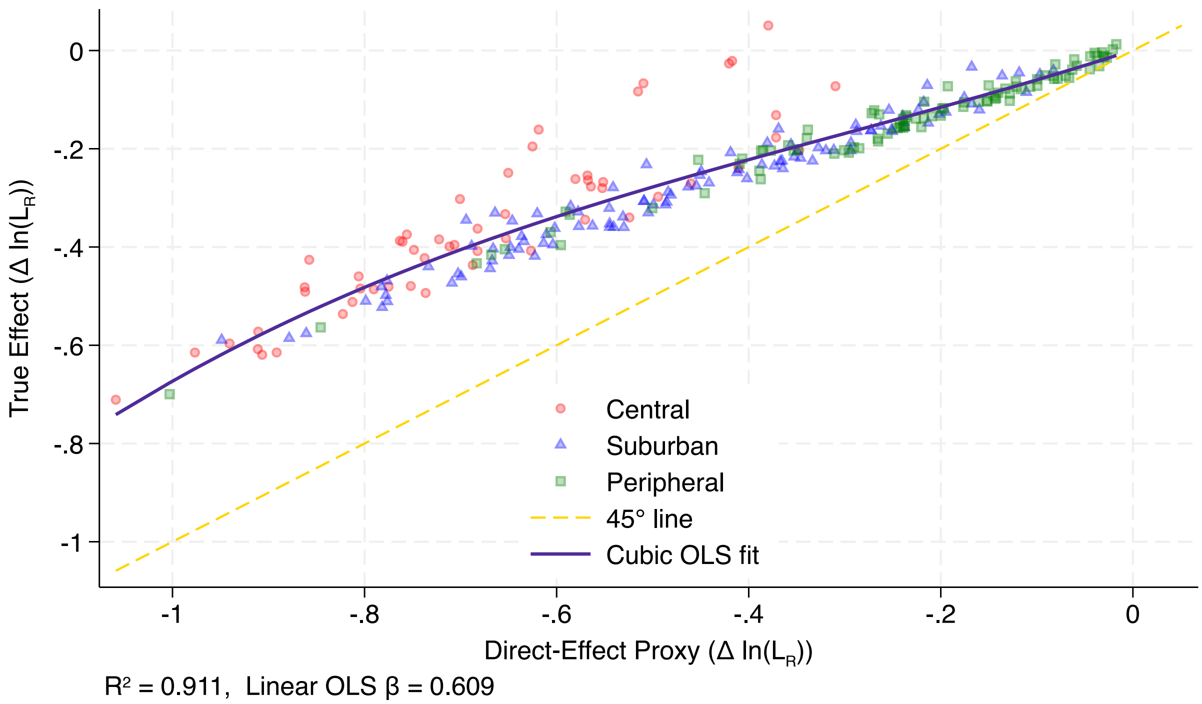


Figure 5: **Exact** General Equilibrium Effects Versus **Exact** Partial Equilibrium Effects of Adding Tokyo's Train System Across Municipalities

References

Gabriel M. Ahlfeldt, Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf. The Economics of Density: Evidence From the Berlin Wall. *Econometrica*, 83(6):2127–2189, 2015. ISSN 1468-0262. doi: 10.3982/ECTA10876. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA10876. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA10876.

Treb Allen and Costas Arkolakis. The Welfare Effects of Transportation Infrastructure Improvements. *The Review of Economic Studies*, 89(6):2911–2957, November 2022. ISSN 0034-6527. doi: 10.1093/restud/rdac001. URL https://doi.org/10.1093/restud/rdac001.

Treb Allen and Costas Arkolakis. Quantitative Regional Economics, January 2025. URL https://papers.ssrn.com/abstract=5122155.

Treb Allen, Costas Arkolakis, and Yuta Takahashi. Universal Gravity. *Journal of Political Economy*, 128(2):393–433, February 2020. ISSN 0022-3808. doi: 10.1086/704385. URL https://www.journals.uchicago.edu/doi/full/10.1086/704385. Publisher: The University of Chicago Press.

Treb Allen, Costas Arkolakis, and Xiangliang Li. On the Equilibrium Properties of Spatial Models. *American Economic Review: Insights*, 6(4):472–489, December 2024. ISSN 2640-205X, 2640-2068. doi: 10.1257/aeri.20230495. URL https://pubs.aeaweb.org/doi/10.1257/aeri.20230495.

Costas Arkolakis, Arnaud Costinot, and Andrés Rodríguez-Clare. New Trade Models, Same Old Gains? *American Economic Review*, 102(1):94–130, February 2012. ISSN 0002-8282. doi: 10.1257/aer.102.1.94. URL https://www.aeaweb.org/articles?id=10.1257/aer.102.1.94.

David Rezza Baqaee and Emmanuel Farhi. Networks, Barriers, and Trade. *Econometrica*, 92(2):505–541, 2024. ISSN 1468-0262. doi: 10.3982/ECTA17513. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA17513. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA17513.

Nathaniel Baum-Snow. Did Highways Cause Suburbanization? *The Quarterly Journal of Economics*, 122(2):775–805, 2007. ISSN 0033-5533. URL https://www.jstor.org/stable/25098858. Publisher: Oxford University Press.

Nathaniel Baum-Snow, Matthew E. Kahn, and Richard Voith. Effects of Urban Rail Transit Expansions: Evidence from Sixteen Cities, 1970-2000 [with Comment]. *Brookings-Wharton Papers on Urban Affairs*, pages 147–206, 2005. ISSN 1528-7084. URL https://www.jstor.org/stable/25067419. Publisher: Brookings Institution Press.

Nathaniel Baum-Snow, Loren Brandt, J. Vernon Henderson, Matthew Turner, and Qinghua Zhang. Roads, Railroads, and Decentralization of Chinese Cities. *The Reveiw of Economics and Statistics*, 99(3):435–448, 2017. URL https://econpapers.repec.org/article/tprrestat/v_3a99_3ay_3a2017_3ai_3a3_3ap_3a435-448.htm. Publisher: MIT Press.

Robert Dekle, Jonathan Eaton, and Samuel Kortum. Global Rebalancing with Gravity: Measuring the Burden of Adjustment. *IMF Staff Papers*, 55(3):511–540, 2008. ISSN 1020-7635. URL https://www.jstor.org/stable/40377786. Publisher: Palgrave Macmillan Journals.

Edward Glaeser, Matthew E. Kahn, and Jordan Rappaport. Why Do The Poor Live In Cities? The Role of Public Transportation. *Journal of Urban Economics*, 2008. ISSN 0094-1190. doi: 10.1016/j.jue.2006.12.004. URL https://dash.harvard.edu/handle/1/2958224. Accepted: 2009-05-14T18:30:49Z Publisher: Elsevier.

Edward L. Glaeser and Matthew E. Kahn. Decentralized Employment and the Transformation of the American City, February 2001. URL https://www.nber.org/papers/w8117.

Edward L. Glaeser and Matthew E. Kahn. Chapter 56 - Sprawl and Urban Growth. In J. Vernon Henderson and Jacques-François Thisse, editors, *Handbook of Regional and Urban Economics*, volume 4 of *Cities and Geography*, pages 2481–2527. Elsevier, January 2004. doi: 10.1016/S1574-0080(04)80013-0. URL https://www.sciencedirect.com/science/article/pii/S1574008004800130.

Marco Gonzalez-Navarro and Matthew A. Turner. Subways and urban growth: Evidence from earth. *Journal of Urban Economics*, 108:85–106, November 2018. ISSN 0094-1190. doi: 10.1016/j.jue.2018.09.002. URL https://www.sciencedirect.com/science/article/pii/S009411901830072X.

Stephan Heblich, Stephen J Redding, and Daniel M Sturm. The Making of the Modern Metropolis: Evidence from London*. *The Quarterly Journal of Economics*, 135(4):2059–2133, November 2020. ISSN 0033-5533. doi: 10.1093/qje/qjaa014. URL https://doi.org/10.1093/qje/qjaa014.

Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985. doi: 10.1017/CBO9780511810817. URL https://www.cambridge.org/core/books/matrix-analysis/9CF2CB491C9E97948B15FAD835EF9A8B.

Benny Kleinman, Ernest Liu, and Stephen J. Redding. International Friends and Enemies. *American Economic Journal: Macroeconomics*, 16(4):350–385, October 2024. ISSN 1945-7707, 1945-7715. doi: 10.1257/mac.20220223. URL https://pubs.aeaweb.org/doi/10.1257/mac.20220223. Publisher: American Economic Association.

Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra, Second Edition*. Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics, January 2023. ISBN 978-1-61197-743-1. doi: 10.1137/1.9781611977448. URL https://epubs.siam.org/doi/book/10.1137/1.9781611977448.

Yuhei Miyauchi, Kentaro Nakajima, and Stephen Redding. The Economics of Spatial Mobility: Theory and Evidence Using Smartphone Data. Technical Report w28497, National Bureau of Economic Research, Cambridge, MA, February 2021. URL http://www.nber.org/papers/w28497.pdf.

Ferdinando Monte, Stephen J. Redding, and Esteban Rossi-Hansberg. Commuting, Migration, and Local Employment Elasticities. *American Economic Review*, 108(12):3855–3890, December 2018. ISSN 0002-8282. doi: 10.1257/aer.20151507. URL https://www.aeaweb.org/articles?id=10.1257/aer.20151507.

Jacques-François Thisse, Matthew A. Turner, and Philip Ushchev. Foundations of cities. *Journal of Urban Economics*, 143:103684, September 2024. ISSN 0094-1190. doi: 10.1016/j.jue.2024.103684. URL https://www.sciencedirect.com/science/article/pii/S0094119024000548.

A Reduced Form Commuting Gravity

In this section, we first introduce three model assumptions such that the model's system of equations can be written with the same structure and the same observables so that our comparative analysis from section 3 is applicable. The purpose of the assumptions is simply to layout the ingredients that imply that the system of equations represent a sum of commuter flows, which are observable. Second, we show in the following section that the reduced form structure can accommodate various assumption on land use provided that land is exogenously separated between commercial and residential use and consumer preferences are quasilinear in the consumption of floorspace.

Suppose there are N locations, indexed $i \in S = \{1, ..., N\}$. Define the residential population, $\{L_{Ri}\} \in \mathbb{R}^N_{++}$ to be the number of workers living in location i, employment $\{L_{Fj}\} \in \mathbb{R}^N_{++}$ as the number of people working in location j. Commuter flows, $\{L_{ij}\} \in \mathbb{R}^{N\times N}_{+}$, is the number of workers who live in i and work in workplace j.

Condition 1. There exists parameters $\phi \in \mathbb{R}_+, \tilde{\alpha} \in \mathbb{R}, \tilde{\beta} \in \mathbb{R}$, exogenous reduced-form fundamentals $\{\tilde{A}_i\} \in \mathbb{R}^N_{++}$, $\{\tilde{U}_i\} \in \mathbb{R}^N_{++}$, and commuting costs $d_{ij} \in \mathbb{R}_{++} \cup \{\infty\}$ for $i, j = 1, ...N$, and an endogenous scalar, $\zeta \in \mathbb{R}_{++}$, such that commuter flows take the form:

$$L_{ij} = \zeta d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j L_{Fj}^{\tilde{\alpha}} L_{Ri}^{\tilde{\beta}} \tag{34}$$

Condition 2. Residential Commuter Market Clearing:

$$\sum_j L_{ij} = L_{Ri} \tag{35}$$

Condition 3. Employment Commuter Market Clearing:

$$\sum_i L_{ij} = L_{Fj} \tag{36}$$

The following theorem establishes the existence of a unique-to-scale solution.

Theorem 7. *Suppose for all $i \in S$, $d_{ij} < \infty$ for at least one j. Similarly for all $j \in S$, $d_{ij} < \infty$ for at least one i. Then there exists a unique-to-scale solution to the system described in Conditions 1, 2, and 3 if* $\frac{|\tilde{\alpha}\tilde{\beta}|}{|(1-\tilde{\beta})(1-\tilde{\alpha})|} < 1$.

In addition, the system can be written in the following "normalized" form:

$$\tilde{L}_{Ri} = \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j \tilde{L}_{Fj}^{\tilde{\alpha}} \tilde{L}_{Ri}^{\tilde{\beta}} \tag{37}$$

$$\tilde{L}_{Fj} = \sum_i d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j \tilde{L}_{Ri}^{\tilde{\beta}} \tilde{L}_{Fj}^{\tilde{\alpha}} \tag{38}$$

where $\tilde{L}_{Ri} \equiv \kappa_R L_{Ri}$ and $\tilde{L}_{Fj} \equiv \kappa_F L_{Fj}$ are the unknowns and κ_R, κ_F are endogenous scalars that depend on population mobility constraints.

Proof. See Section B □

Given our assumptions, residential commuter shares, $\frac{\mathbf{L}}{\mathbf{L_R}}$ and workplace commuter shares $\frac{\mathbf{L^T}}{\mathbf{L_F}}$ form our "flow share" expressions from theorem 1. The remaining results from section 3 follow directly, once we substitute $\alpha\theta = \tilde{\alpha}, \quad \beta\theta = \tilde{\beta}, \quad \theta = \phi$

A.1 Isomorphisms

We show in this section that the reduced form structure in section A can accomodate various assumptions on land use provided that land is exogenously separated between commercial and residential use and consumer preferences are quasilinear in the consumption of floorspace.

Derivations

A.2 Deriving Equilibrium System of Equations

Using the commuter flow equation (eq. 21) and residential and workplace market clearing conditions (eq. 22 and23) and making the substitution

$$\kappa \equiv \frac{\sum_i \sum_j (d_{ij} \bar{U}_i L_{Ri}^{\beta} \bar{A}_j L_{Fj}^{\alpha})^{\theta}}{\bar{L}} \tag{39}$$

we have the following system of $2 \times N + 1$ equations and $2 \times N + 1$ unknowns.

$$L_{Ri} = \kappa^{-1} \sum_j d_{ij}^{-\theta} \bar{U}_i^\theta (L_{Ri})^{\beta\theta} \bar{A}_j^\theta (L_{Fj})^{\alpha\theta} \tag{40}$$

$$L_{Fj} = \kappa^{-1} \sum_i d_{ij}^{-\theta} \bar{U}_i^\theta (L_{Ri})^{\beta\theta} \bar{A}_j^\theta (L_{Fj})^{\alpha\theta} \tag{41}$$

$$\kappa = \frac{\sum_i \sum_j (d_{ij}^{-1} \bar{U}_i L_{Ri}^\beta \bar{A}_j L_{Fj}^\alpha)^\theta}{\bar{L}} \tag{42}$$

These equations satisfy conditions 1, 2, and 3 in theorem 7, where $\tilde{U}_i = \bar{U}_i^\theta, \tilde{A}_i = \bar{A}_i^\theta, \phi = \theta, \tilde{\beta} = \theta\beta, \tilde{\alpha} = \alpha\beta, \zeta = \kappa^{-1}$.

Thus if the following conditions hold:

1. $d_{ij} \in \mathbb{R}_{++} \cup \{\infty\}$,
2. For all $i \in S$, $d_{ij} < \infty$ for at least one j,
3. For all $j \in S$, $d_{ij} < \infty$ for at least one i,
4. $\{\tilde{A}_i\} \in \mathbb{R}_{++}^N$, $\{\tilde{U}_i\} \in \mathbb{R}_{++}^N$
5. $\theta \in \mathbb{R}_+$
6. $\frac{|\alpha\theta\beta\theta|}{|(1-\beta\theta)(1-\alpha\theta)|} < 1$

, then there exists a unique-to-scale solution with $\{L_{Ri}\} \in \mathbb{R}_{++}^N$, $\{L_{Fj}\} \in \mathbb{R}_{++}^N$, to the system.

Theorem 7 uses the results from Allen et al. (2024) to establish existence of a unique solution. The authors also offer a method to rid the system of quations of endogenous scalars such as κ by absorbing them into the redefined endogenous variables. We first rearrange the first two equations:

$$L_{Ri}^{1-\beta\theta} = \kappa^{-1} \sum_j d_{ij}^{-\theta} \bar{U}_i^\theta \bar{A}_j^\theta (L_{Fj})^{\alpha\theta}$$

$$L_{Fj}^{1-\alpha\theta} = \kappa^{-1} \sum_i d_{ij}^{-\theta} \bar{U}_i^\theta \bar{A}_j^\theta (L_{Rj})^{\beta\theta}$$

and define the change of variables:

$$y_{Ri} \equiv L_{Ri}^{1-\beta\theta} \kappa^{D_R}$$

$$y_{Fj} \equiv L_{Fj}^{1-\alpha\theta} \kappa^{D_F}$$

The first two equations of the system can then be rewritten as:

$$\kappa^{1-D_R} y_{Ri} = \sum_j d_{ij}^{-\theta} \bar{U}_i^{\theta} \bar{A}_j^{\theta} (y_{Fj})^{\frac{\alpha\theta}{1-\alpha\theta}} \kappa^{-D_F \frac{\alpha\theta}{1-\alpha\theta}} \tag{43}$$

$$\kappa^{1-D_F} y_{Fj} = \sum_i d_{ij}^{-\theta} \bar{U}_i^{\theta} \bar{A}_j^{\theta} (y_{Ri})^{\frac{\beta\theta}{1-\beta\theta}} \kappa^{-D_F \frac{\beta\theta}{1-\beta\theta}} \tag{44}$$

To eliminate the dependence on κ, we set $\mathbf{D} = [D_R \quad D_F]'$ as the solution to the linear system:

$$\begin{bmatrix} 1 \\ \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{-\alpha\theta}{1-\alpha\theta} \\ \\ \frac{-\beta\theta}{1-\beta\theta} & 1 \end{bmatrix} \begin{bmatrix} D_R \\ \\ D_F \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} D_R \\ \\ D_F \end{bmatrix} = \begin{bmatrix} 1 & \frac{-\alpha\theta}{1-\alpha\theta} \\ \\ \frac{-\beta\theta}{1-\beta\theta} & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ \\ 1 \end{bmatrix}$$

We thus have the following system of $2 \times N$ equations of $2 \times N$ unknowns, where κ is no longer present in the equations:

$$y_{Ri} = \sum_j d_{ij}^{-\theta} \bar{U}_i^{\theta} \bar{A}_j^{\theta} y_{Fj}^{\frac{\alpha\theta}{1-\alpha\theta}}$$

$$y_{Fj} = \sum_i d_{ij}^{-\theta} \bar{U}_i^{\theta} \bar{A}_j^{\theta} y_{Ri}^{\frac{\beta\theta}{1-\beta\theta}}$$

Substituting back the original variables, and now defining $\tilde{L}_{Ri} \equiv \kappa_R L_{Ri}$ and $\tilde{L}_{Fj} \equiv \kappa_F L_{Fj}$ where $\kappa_R \equiv \kappa^{\frac{D_R}{1-\beta\theta}}$ and $\kappa_F \equiv \kappa^{\frac{D_F}{1-\alpha\theta}}$, we can rewrite the system as:

$$\tilde{L}_{Ri} = \sum_j d_{ij}^{-\theta} \bar{U}_i^\theta \bar{A}_j^\theta \tilde{L}_{Fj}^{\alpha\theta} \tilde{L}_{Ri}^{\beta\theta}$$

$$\tilde{L}_{Fj} = \sum_i d_{ij}^{-\theta} \bar{U}_i^\theta \bar{A}_j^\theta \tilde{L}_{Ri}^{\beta\theta} \tilde{L}_{Fj}^{\alpha\theta}$$

as desired.

The population and employment variables, $\tilde{L}_{Ri}$ and $\tilde{L}_{Fj}$, are now expressed as a product of the true population and employment variables, L_{Ri} and L_{Fj}, and the endogenous scalars κ_R and κ_F, which depend on the population-mobility assumptions. These equations are therefore useful for studying relative changes in population (e.g. L_{R1}/L_{R2}) and employment (e.g. L_{F1}/L_{F2}).

B Proofs

B.1 Proof of Theorem 1

Proof. Denote $\boldsymbol{\gamma} \equiv (\boldsymbol{\Gamma} - \boldsymbol{R})$. Note also that we defined $\boldsymbol{\alpha} \equiv \mathbf{B}(\boldsymbol{\Gamma} - \mathbf{R})^{-1} = \mathbf{B}\boldsymbol{\gamma}^{-1}$

Part 1. is a consequence of the implicit function theorem. We use the theorem to first derive comparative statics for composite variables, $\frac{\partial \ln \mathbf{y}^*}{\partial \ln \boldsymbol{\theta}} = \bar{\mathbf{A}}^{-1}\mathbf{T}$. Then given the change of variables relationship: $\ln \mathbf{y}^* = \left(\boldsymbol{\gamma} \otimes \mathbf{I}_N\right) \ln \mathbf{x}^*$, we have that $\frac{\partial \ln \mathbf{x}^*}{\partial \ln \boldsymbol{\theta}} = \left(\boldsymbol{\gamma} \otimes \mathbf{I}_N\right)^{-1} \frac{\partial \ln \mathbf{y}^*}{\partial \ln \boldsymbol{\theta}} = \left(\boldsymbol{\gamma} \otimes \mathbf{I}_N\right)^{-1} \bar{\mathbf{A}}^{-1}\mathbf{T} = \left(\bar{\mathbf{A}}\left(\boldsymbol{\gamma} \otimes \mathbf{I}_N\right)\right)^{-1}\mathbf{T}$.

It suffices then to show that $\frac{\partial \ln \mathbf{y}^*}{\partial \ln \boldsymbol{\theta}} = \bar{\mathbf{A}}^{-1}\mathbf{T}$, using the implicit function theorem.

Given the transformed system (7), taking logs of the left and right and subtracting, we define the function $F_{ih} : \mathbb{R}^{NH} \times \Theta \to \mathbb{R}^{NH}$ as follows:

$$F_{i,h}(\mathbf{y}, \boldsymbol{\theta}) = \ln\left(y_{ih}\right) - \ln\left(\sum_{j \in \mathcal{N}} K_{ijh}(\boldsymbol{\theta}) \prod_{h' \in \mathcal{H}} y_{jh'}^{\alpha_{hh'}}\right)$$

for $i = 1, \ldots, N, h = 1, \ldots, H$. We stack the functions as $\boldsymbol{F} = [F_{11}, \ldots, F_{N1}, F_{12}, \ldots, F_{N2}, F_{13}, \ldots, F_{NH}]'$. The implicit function theorem is applied to this vector of equations by taking the partial elasticity with respect to the block vector $\mathbf{y} \equiv [y_{11}, \ldots, y_{N,1} y_{12}, \ldots, y_{N2}, y_{13}, \ldots, y_{NH}]'$. Taking the partial elasticity of $\frac{\partial \ln(y_{ih'})}{\partial \ln yjh'} = 1$ if $i = j, h = h'$. In matrix form, $\frac{\partial \ln(\mathbf{y})}{\partial \ln \mathbf{y'}} = \mathbf{I}_{HN}$. For the second term, we have: $\frac{\partial \ln\left(\sum_{j \in \mathcal{N}} K_{ijh}(\boldsymbol{\theta}) \prod_{h' \in \mathcal{H}} y_{jh'}^{\alpha_{hh'}}\right)}{\partial \ln \mathbf{y}}$ for $i = 1, \ldots, N; h = 1, \ldots, H$.

When taking the derivative of with i, h-th term with respect to $\ln(y_{jh'})$, we have:

$$\frac{\partial \ln\left(\sum_{j\in\mathcal{N}} K_{ijh}(\boldsymbol{\theta})\prod_{h'\in\mathcal{H}} y_{jh'}^{\alpha_{hh'}}\right)}{\partial \ln y_{jh'}} = \alpha_{hh'}\frac{K_{ijh}\prod_{h'\in\mathcal{H}} y_{jh'}^{\alpha_{hh'}}}{\sum_{j'} K_{ij'h}\prod_{h'\in\mathcal{H}} y_{j'h'}^{\alpha_{hh'}}} = \alpha_{hh'}\frac{K_{ijh}\prod_{h'\in\mathcal{H}} y_{jh'}^{\alpha_{hh'}}}{y_{ih}}$$

We note that fractional term $\left[\mathbf{X}_{hh'}\right]_{ij} \equiv \frac{K_{ijh}\prod_{h'\in\mathcal{H}} y_{jh'}^{\alpha_{hh'}}}{y_{ih}}$ does not depend on h', such that we write $\left[\mathbf{X}_{hh'}\right]_{ij} = \left[\mathbf{X}_{(h)}\right]_{ij}$. We note that $\mathbf{X}_{(h)}$ is a row-stochastic matrix. The elements of a row sum to one.

Putting together the coefficient vector, we have that each block, h, h' of the $NH \times NH$ matrix is equal to $\alpha_{hh'}\left[\mathbf{X}_{(h)}\right]$ This can be written as $\operatorname{bdiag}\left(\mathbf{X}_{(1)}, \ldots, \mathbf{X}_{(H)}\right)(\boldsymbol{\alpha}\otimes \boldsymbol{I}_N)$, which we denote as $\boldsymbol{\alpha}_X$. Thus, we have $\bar{A} = \mathbf{I}_{HN} - \boldsymbol{\alpha}_X$.

Taking the derivative with respect to $\ln\boldsymbol{\theta'}$ results in dependence on $\mathbf{X}_{(h)}$ again, leading to the block-stacked matrix:$\mathbf{T}\in\mathbb{R}^{HN\times M}$ where $\mathbf{T} = \begin{bmatrix}\mathbf{T}_1' & \cdots & \mathbf{T}_H'\end{bmatrix}'$ with $\mathbf{T}_h \in \mathbb{R}^{N\times M}$ with entries:

$$\left[\mathbf{T}_h\right]_i = \sum_{j=1}^{N}\left[\mathbf{X}_{(h)}\right]_{ij}\frac{\partial \ln K_{ijh}}{\partial \ln\boldsymbol{\theta'}} \qquad (h = 1,\ldots,H;\ i = 1,\ldots,N).$$

Part 2.

We have that $\overline{\mathbf{A}} \equiv \mathbf{I}_{HN} - \boldsymbol{\alpha}_X$. Then $\overline{\mathbf{A}}^{-1}$ can be written as the following Neumann series if and only if $\rho(\alpha_X) < 1$. $\overline{\mathbf{A}}^{-1} = \sum_{k=0}^{\infty}\boldsymbol{\alpha}_X^k$. See theorems 7.10.8-11 of Meyer (2023).

We show that $\|\boldsymbol{\alpha}_X\|_\infty = \||\boldsymbol{\alpha}_X|\|_\infty = \|\boldsymbol{\alpha}\|_\infty = \||\boldsymbol{\alpha}|\|_\infty$. For any matrix A, the absolute value does not change the value of the norm. Thus, it suffices to show $\|\boldsymbol{\alpha}_X\|_\infty = \|\boldsymbol{\alpha}\|_\infty$. We note that $|\operatorname{bdiag}\left(\mathbf{X}_{(1)},\ldots,\mathbf{X}_{(H)}\right)(\boldsymbol{\alpha}\otimes\boldsymbol{I}_N)| = \operatorname{bdiag}\left(\mathbf{X}_{(1)},\ldots,\mathbf{X}_{(H)}\right)|(\boldsymbol{\alpha}\otimes\boldsymbol{I}_N)| = \operatorname{bdiag}\left(\mathbf{X}_{(1)},\ldots,\mathbf{X}_{(H)}\right)(|\boldsymbol{\alpha}|\otimes\boldsymbol{I}_N)$. The first equality is due to the non-negativity of $\mathbf{X}$. $\operatorname{bdiag}\left(\mathbf{X}_{(1)},\ldots,\mathbf{X}_{(H)}\right)(|\boldsymbol{\alpha}|\otimes\boldsymbol{I}_N)$ is a block matrix where each block h, h' is equal to $|\boldsymbol{\alpha_{hh'}}|\mathbf{X}_{(h)}$ Because $\mathbf{X}_{(h)}$ is row stochastic, the row sum of each block $|\boldsymbol{\alpha_{hh'}}|\mathbf{X}_{(h)}$ is $|\boldsymbol{\alpha_{hh'}}|$, and thus the row sum across all blocks is equal to $\sum_{h'}|\boldsymbol{\alpha_{hh'}}|$, the same as the row sum of row h' of $\boldsymbol{\alpha}$. Thus, the maximum absolute row sums are the same between the two.

We demonstrate the properties of $\rho(\boldsymbol{\alpha}_X)$. For any matrix,A, $\rho(A)\le\rho(|A|)$. (See for example Theorem 8.1.18. of Horn and Johnson (1985)). Thus $\rho(\boldsymbol{\alpha}_X)\le\rho(|\boldsymbol{\alpha}_X|)$.

We next show that $\rho(\boldsymbol{\alpha}_X)\ge\rho(\boldsymbol{\alpha})$.

We have that $\boldsymbol{\alpha}_X \equiv \operatorname{bdiag}\left(\mathbf{X}_{(1)},\ldots,\mathbf{X}_{(H)}\right)(\boldsymbol{\alpha}\otimes\boldsymbol{I}_N)$. For any vector $\boldsymbol{\xi}\in\mathbb{R}^H$, let $\tilde{\boldsymbol{\xi}}$ be the block-constant vector $\tilde{\boldsymbol{\xi}} \equiv \boldsymbol{\xi} \otimes \mathbf{1}_N$. For any vector $\boldsymbol{\xi} \in \mathbb{R}^H$, we then have:

$$(\boldsymbol{\alpha} \otimes \boldsymbol{I}_N)\tilde{\boldsymbol{\xi}} = (\boldsymbol{\alpha}\boldsymbol{\xi}) \otimes \mathbf{1}_N$$
$$\text{diag}\Big(\mathbf{X}_{(1)}, \ldots, \mathbf{X}_{(H)}\Big)\tilde{\boldsymbol{\xi}} = \tilde{\boldsymbol{\xi}} = \boldsymbol{\xi} \otimes \mathbf{1}_N$$

where the second line follows from the property that each block $\mathbf{X}_{(h)}$ is row-stochastic for all h and therefore $\mathbf{X}_{(h)}\mathbf{1}_N = \mathbf{1}_N$ for all h.

Applying the matrices in order, we thus have:

$$\boldsymbol{\alpha}_X\tilde{\xi} = \text{bdiag}\Big(\mathbf{X}_{(1)}, \ldots, \mathbf{X}_{(H)}\Big)\,(\boldsymbol{\alpha} \otimes \boldsymbol{I}_N)\tilde{\xi} = (\boldsymbol{\alpha}\boldsymbol{\xi}) \otimes \mathbf{1}_N$$

Hence for any $\boldsymbol{v} \in \mathbb{R}^H$ where $\boldsymbol{\alpha}\boldsymbol{v} = \lambda\boldsymbol{v}$, then $\boldsymbol{\alpha}_X\,(\boldsymbol{v} \otimes \mathbf{1}_N) = \lambda\,(\boldsymbol{v} \otimes \mathbf{1}_N)$. In other words, any eigenvalue of $\boldsymbol{\alpha}$ is also an eigenvalue of $\boldsymbol{\alpha}_X$. Thus, we have that $\rho(\boldsymbol{\alpha}_X) \geq \rho(\boldsymbol{\alpha})$.

Next, we show that $\rho(|\boldsymbol{\alpha}_X|) = \rho(|\boldsymbol{\alpha}|)$. The same argument as above can be used to show that $\rho(|\boldsymbol{\alpha}_X|) \geq \rho(|\boldsymbol{\alpha}|)$.

Collatz-Wieland's formula states that $\rho(|\boldsymbol{\alpha}_X|) \leq \max_i \frac{(|\boldsymbol{\alpha}_X|\boldsymbol{v})_i}{\boldsymbol{v}_i}$ for any positive vector $\boldsymbol{v}$. The nonnegativity of $|\boldsymbol{\alpha}|$ implies that there exists a nonnegative $\boldsymbol{w}$ of $|\boldsymbol{\alpha}|$ for which the eigenvalue is $\lambda = \rho(|\boldsymbol{\alpha}|) > 0$ (see for example Theorem 8.3.1 in Horn and Johnson (1985)). Let $\epsilon > 0$, $\boldsymbol{y} = \boldsymbol{w} + \epsilon\mathbf{1}_H$ such that $\boldsymbol{y}$ is positive. Then choose $\boldsymbol{v} = \boldsymbol{y} \otimes \mathbf{1}_N$. We have that $\max_i \frac{(|\boldsymbol{\alpha}_X|\boldsymbol{v})_i}{\boldsymbol{v}_i} = \rho(|\boldsymbol{\alpha}|) + \max_i \frac{(|\boldsymbol{\alpha}|\epsilon\mathbf{1}_H)_i}{\boldsymbol{y}_i}$, whose limit as $\epsilon \to 0$ implies $\rho(|\boldsymbol{\alpha}_X|) \leq \rho(|\boldsymbol{\alpha}|)$. Combining both results, we have $\rho(|\boldsymbol{\alpha}_X|) = \rho(|\boldsymbol{\alpha}|)$.

ii.

The zeroth term of the comparative static is equal to $-\tilde{\mathbf{\Gamma}}^{-1}\mathbf{T}$, while the full general equilibrium effect is equal to $-\tilde{\mathbf{\Gamma}}^{-1}\sum_{k=0}^{\infty}\boldsymbol{\alpha}_X^k\mathbf{T}$.

Taking the partial derivative of $\tilde{g}(\cdot, K_{ijh}(\theta))$ results in: $\sum_j \frac{\partial \ln g_{ih}}{\partial \ln K_{ijh}}\frac{\partial \ln K_{ijh}}{\partial \ln \kappa} = \sum_j X_{ijh}\frac{\partial \ln K_{ijh}}{\partial \ln \kappa}$, resulting in the same structure as equation 10. This is equal to the local partial equilibrium effect on composite variable y_{ih}. In matrix notation, we have $\mathbf{b} = \mathbf{T}$, so we have $\tilde{\mathbf{\Gamma}}^{-1}\mathbf{T} = \tilde{\mathbf{\Gamma}}^{-1}\mathbf{b}$ as desired. If we are considering a vector, then we use the row vector $\boldsymbol{\theta}'$ instead of κ.

Part iii.

For any norm $\|\cdot\|$ and dual norm $\|\cdot\|^D$, and all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$, $|\mathbf{y}'\mathbf{x}| \leq \|\mathbf{x}\|\|\mathbf{y}\|^D$ and $\|\mathbf{y}'\mathbf{x}| \leq \|\mathbf{x}\|^D\|\mathbf{y}\|$. See Lemma 5.4.13 in Horn and Johnson (1985)

For vectors $u \in \mathbb{R}^{HN} \quad v \in \mathbb{R}^M$ and choosing the 1-norm and it's dual, the $\infty-$ norm, we have: $|u'\mathbf{R^1}\mathbf{T}v| \leq \|u\|_1\|\mathbf{R^1}\mathbf{T}v\|_\infty \leq \|u\|_1\|\mathbf{R^1}\|_\infty\|\mathbf{T}v\|_\infty \leq \|u\|_1\|\tilde{\mathbf{\Gamma}}^{-1}\|_\infty\|\mathbf{Tv}\|_\infty\frac{\|\boldsymbol{\alpha}_{\boldsymbol{X}}\|_\infty}{1-\|\boldsymbol{\alpha}_{\boldsymbol{X}}\|_\infty} \leq$

$\|u\|_1 \|\tilde{\mathbf{\Gamma}}^{-1}\|_\infty \|\mathbf{T}\|_\infty \|\mathbf{v}\|_\infty \frac{\|\boldsymbol{\alpha}_{\boldsymbol{X}}\|_\infty}{1-\|\boldsymbol{\alpha}_{\boldsymbol{X}}\|_\infty}$.
Note that one can also write:
$|u'\mathbf{R^1}\mathbf{T}v| \leq \|(u\mathbf{\Gamma}^{-1})'\|_1 \| \sum_{k=1} \boldsymbol{\alpha}_{\boldsymbol{X}}^k \mathbf{T} v\|_\infty \leq \|(u\mathbf{\Gamma}^{-1})'\|_1 \|\mathbf{T}v\|_\infty \frac{\|\boldsymbol{\alpha}_{\boldsymbol{X}}\|_\infty}{1-\|\boldsymbol{\alpha}_{\boldsymbol{X}}\|_\infty}$

□

B.2 Proof of Theorem 2

Proof. Given $\mathbf{V} \in \mathbb{R}^{H \times H}$, multiplying both sides of Equation (5) by $\prod_{h' \in \mathcal{H}} x_{ih}^{V_{hh'}}$ yields

$$\prod_{h' \in \mathcal{H}} x_{ih}^{\gamma_{hh'} - \rho_{hh'}} \prod_{h' \in \mathcal{H}} x_{ih}^{V_{hh'}} = \prod_{h' \in \mathcal{H}} x_{ih}^{\gamma_{hh'} - \rho_{hh'} + V_{hh'}} = \prod_{h' \in \mathcal{H}} x_{ih}^{V_{hh'}} \sum_{j \in \mathcal{N}} K_{ijh} \prod_{h' \in \mathcal{H}} x_{jh'}^{\beta_{hh'}}.$$

Applying the change of variables $y_{ih} \equiv \prod_{h' \in \mathcal{H}} x_{ih}^{\gamma_{hh'} - \rho_{hh'} + V_{hh'}}$, we obtain

$$y_{ih} = \prod_{h' \in \mathcal{H}} y_{ih'}^{\tilde{V}_{hh'}} \sum_{j \in \mathcal{N}} K_{ijh} \prod_{h' \in \mathcal{H}} y_{jh'}^{\breve{\alpha}_{hh'}},$$

where $\tilde{V}_{hh'}$ are the (h, h') elements of $\tilde{\mathbf{V}} \equiv \mathbf{V}\breve{\boldsymbol{\gamma}}^{-1}$ and $\breve{\alpha}_{hh'}$ are the (h, h') elements of $\breve{\boldsymbol{\alpha}} \equiv \mathbf{B}\breve{\boldsymbol{\gamma}}^{-1}$. Applying the implicit function theorem as in Theorem 1 yields the same $\mathbf{T}$ and $\mathbf{A}$, but with $\mathbf{A} \equiv \bar{\breve{\mathbf{A}}}\,\breve{\breve{\mathbf{\Gamma}}}$, where

$$\breve{\breve{\mathbf{\Gamma}}} \equiv \breve{\boldsymbol{\gamma}} \otimes \mathbf{I}_N, \qquad \bar{\breve{\mathbf{A}}} \equiv \mathbf{I}_{HN} - \breve{\boldsymbol{\alpha}}_X, \qquad \breve{\boldsymbol{\alpha}}_X(\mathbf{V}) \equiv (\tilde{\mathbf{V}} \otimes \mathbf{I}_N) + \text{bdiag}\Big(\mathbf{X}_{(1)}, \ldots, \mathbf{X}_{(H)}\Big)(\breve{\boldsymbol{\alpha}} \otimes \mathbf{I}_N).$$

Part 2. When we choose $\mathbf{V} = \boldsymbol{\gamma} - \text{diag}(\boldsymbol{\gamma})$, the zeroth-degree term is $\breve{\breve{\mathbf{\Gamma}}}^{-1}\mathbf{T} = (\breve{\boldsymbol{\gamma}}^{-1} \otimes \mathbf{I}_N)\mathbf{T}$. Because the diagonal matrix $\breve{\boldsymbol{\gamma}}^{-1}$ has elements $\frac{1}{\gamma_{hh} - \rho_{hh}}$, the (i, h) element of the zeroth-degree effect is

$$[(\breve{\boldsymbol{\gamma}}^{-1} \otimes \mathbf{I}_N)\mathbf{T}]_{ih} = \frac{1}{\gamma_{hh} - \rho_{hh}} \sum_{j=1}^{N} \Big[\mathbf{X}_{(h)}\Big]_{ij} \frac{\partial \ln K_{ijh}}{\partial \ln \boldsymbol{\theta'}}.$$

Similarly, for the single-variable local partial-equilibrium effect we have

$$\frac{\partial^{SLP} \ln x_{ih}}{\partial \ln \kappa} \equiv \sum_j \frac{\partial \ln f_{ih}}{\partial \ln K_{ijh}} \frac{\partial \ln K_{ijh}}{\partial \ln \kappa}.$$

Taking the derivative of $f_{ih}(\cdot)$ gives $\frac{\partial \ln f_{ih}}{\partial \ln K_{ijh}} = \frac{1}{\gamma_{hh} - \rho_{hh}} \cdot X_{ijh}$. Thus the full partial-equilibrium effect is

$$\sum_j \frac{\partial \ln f_{ih}}{\partial \ln K_{ijh}} \frac{\partial \ln K_{ijh}}{\partial \ln \kappa} = \sum_j \frac{1}{\gamma_{hh} - \rho_{hh}} \cdot \Big[\mathbf{X}_{(h)}\Big]_{ij} \frac{\partial \ln K_{ijh}}{\partial \ln \kappa},$$

as desired. □

B.3 Proof of Theorem in Section B.3

Proposition. If $\rho(\boldsymbol{\alpha}) = \rho(|\boldsymbol{\alpha}|)$, then $\rho(\boldsymbol{\alpha}_X) = \rho(|\boldsymbol{\alpha}_X|) = \rho(\boldsymbol{\alpha}) = \rho(|\boldsymbol{\alpha}|)$.

Proof. We have $\rho(\boldsymbol{\alpha}) \leq \rho(\boldsymbol{\alpha}_X) \leq \rho(|\boldsymbol{\alpha}_X|) = \rho(|\boldsymbol{\alpha}|)$. If $\rho(\boldsymbol{\alpha}) = \rho(|\boldsymbol{\alpha}|)$ the inequality collapses. □

Definition. A matrix $\mathbf{A} \in \mathbb{R}^{H \times H}$ is *sign-similar to a non-negative matrix* if there exists a matrix $\boldsymbol{\Sigma} \equiv \text{diag}\Big(\sigma_1, \ldots, \sigma_{(H)}\Big)$ with each $\sigma_i \in \{+1, -1\}$, such that $\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma} \geq 0$

One interpretation of sign-similarity is if we viewed the signs of $\boldsymbol{\alpha}$ as the adjacency matrix of a directed graph with signs. The graph represents a matrix that is sign-similar to a non-negative matrix if every cycle contains an even number of negative edges.

Proposition. Sign-Similarity

If $\boldsymbol{\alpha}$ is sign-similar to a non-negative matrix, then $\rho(\boldsymbol{\alpha}) = \rho(\boldsymbol{\alpha}_X) = \rho(|\boldsymbol{\alpha}|) = \rho(|\boldsymbol{\alpha}_X|)$.

Proof. The spectral radius is invariant under conjugation: given $\boldsymbol{\Sigma}$, $\rho(\boldsymbol{\Sigma}\alpha\boldsymbol{\Sigma}^{-1}) = \rho(\alpha)$. $\boldsymbol{\Sigma}\boldsymbol{\alpha}\boldsymbol{\Sigma}^{-1} = |\boldsymbol{\alpha}|$ □

B.4 Proof of Theorem 7

Proof. Substituting the commuter flow equation (34) into residential and workplace market clearing conditions, (36) and (35), yields the following $2 \times N$ equations and $2 \times N + 1$ unknowns.

$$L_{Ri} = \zeta \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j L_{Fj}^{\tilde{\alpha}} L_{Ri}^{\tilde{\beta}} \tag{45}$$

$$L_{Fj} = \zeta \sum_i d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j L_{Fj}^{\tilde{\alpha}} L_{Ri}^{\tilde{\beta}} \tag{46}$$

Allen et al. (2024) characterize the equilibrium existence and uniqueness properties of systems of the above type. In this case, we define the change of variables:

$$y_{Ri} \equiv L_{Ri}^{1-\tilde{\beta}} \zeta^{\tilde{D}_R} \tag{47}$$

$$y_{Fj} \equiv L_{Fj}^{1-\tilde{\alpha}} \zeta^{\tilde{D}_F} \tag{48}$$

where

$$\begin{bmatrix} \tilde{D}_R \\ \\ \tilde{D}_F \end{bmatrix} = \begin{bmatrix} 1 & \frac{-\tilde{\alpha}}{1-\tilde{\alpha}} \\ \\ \frac{-\tilde{\beta}}{1-\tilde{\beta}} & 1 \end{bmatrix}^{-1} \begin{bmatrix} -1 \\ \\ -1 \end{bmatrix}$$

which is well defined if $\tilde{\beta} \neq 1, \tilde{\alpha} \neq 1$ and if the determinate of the matrix $\begin{bmatrix} 1 & \frac{-\tilde{\alpha}}{1-\tilde{\alpha}} \\ \\ \frac{-\tilde{\beta}}{1-\tilde{\beta}} & 1 \end{bmatrix}$ is non-zero: $1 - \frac{\tilde{\alpha}}{1-\tilde{\alpha}}\frac{\tilde{\beta}}{1-\tilde{\beta}} \neq 0$, or equivalently, that $\tilde{\alpha} + \tilde{\beta} \neq 1$ when $\tilde{\alpha} \neq 1, \quad \tilde{\beta} \neq 1$, which our hypothesis satisifes.

Substitution from the change-of-variables yields the following system of $2 \times N$ equations of $2 \times N$ unknowns, where ζ is no longer present in the equations:

$$y_{Ri} = \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Fj}^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}}} \tag{49}$$

$$y_{Fj} = \sum_i d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Ri}^{\frac{\tilde{\beta}}{1-\tilde{\beta}}} \tag{50}$$

The hypothesis of Theorem 1 (i) in Allen et al. (2024) with Remark 1, stated below in the context of the present model, concerns the bounds to the following expressions when $y_{Ri} \in \mathbb{R}_{++}, y_{Fi} \in \mathbb{R}_{++}$: $\sum_k \left| \frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Fj}^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}}}}{\partial \ln y_{Rk}} \right|$, $\sum_k \left| \frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Fj}^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}}}}{\partial \ln y_{Fk}} \right|$, $\sum_k \left| \frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Ri}^{\frac{\tilde{\beta}}{1-\tilde{\beta}}}}{\partial \ln y_{Rk}} \right|$, $\sum_k \left| \frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Ri}^{\frac{\tilde{\beta}}{1-\tilde{\beta}}}}{\partial \ln y_{Fk}} \right|$.

We have: $\frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Fj}^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}}}}{\partial \ln y_{Rk}} = 0$, $\frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Fj}^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}}}}{\partial \ln y_{Fk}} = \frac{\tilde{\alpha}}{1-\tilde{\alpha}} \frac{d_{ik}^{-\phi} \tilde{U}_i \tilde{A}_k y_{Fk}^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}}}}{\sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Fj}^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}}}}$, whose denominator is well defined if for all i, $d_{ij} < \infty$ for at least one j, given that the variables $\tilde{U}_i, \tilde{A}_j, y_{Fj}$ are strictly positive. Thus, $\sum_k \left| \frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Fj}^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}}}}{\partial \ln y_{Fk}} \right| = \left| \frac{\tilde{\alpha}}{1-\tilde{\alpha}} \right|$. Similarly, $\sum_k \left| \frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Ri}^{\frac{\tilde{\beta}}{1-\tilde{\beta}}}}{\partial \ln y_{Rk}} \right| = \left| \frac{\tilde{\beta}}{1-\tilde{\beta}} \right|$ if for all j, $d_{ij} < \infty$ for at least one i, and $\sum_k \left| \frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Ri}^{\frac{\tilde{\beta}}{1-\tilde{\beta}}}}{\partial \ln y_{Fk}} \right| = 0$

The hypothesis of Theorem 1 (i) in Allen et al. (2024) with Remark 1 requires that the spectral radius of the matrix of bounds: $\begin{bmatrix} 0 & \left|\frac{\tilde{\alpha}}{1-\tilde{\alpha}}\right| \\ \left|\frac{\tilde{\beta}}{1-\tilde{\beta}}\right| & 0 \end{bmatrix}$, is less than one. The largest eigenvalue of the matrix can be shown to be: $\rho(\begin{bmatrix} 0 & \left|\frac{\tilde{\alpha}}{1-\tilde{\alpha}}\right| \\ \left|\frac{\tilde{\beta}}{1-\tilde{\beta}}\right| & 0 \end{bmatrix}) = \sqrt{\left|\frac{\tilde{\alpha}}{1-\tilde{\alpha}}\right| \left|\frac{\tilde{\beta}}{1-\tilde{\beta}}\right|}$.

Our hypothesis states that $\left|\frac{\tilde{\alpha}}{1-\tilde{\alpha}}\right| \left|\frac{\tilde{\beta}}{1-\tilde{\beta}}\right|$, which implies that $\sqrt{\left|\frac{\tilde{\alpha}}{1-\tilde{\alpha}}\right| \left|\frac{\tilde{\beta}}{1-\tilde{\beta}}\right|} < 1$.

Thus by Theorem 1 (i) in Allen et al. (2024) with Remark 1, there exists a unique solution to the system (49) and (50), implying that there exists a unique-to-scale solution to the system (45) and (46).

Substituting back the original variables, and now defining $\tilde{L}_{Ri} \equiv \kappa_R L_{Ri}$ and $\tilde{L}_{Fj} \equiv \kappa_F L_{Fj}$ where $\kappa_R \equiv \zeta^{\frac{\tilde{D}_R}{1-\tilde{\beta}}}$ and $\kappa_F \equiv \zeta^{\frac{\tilde{D}_F}{1-\tilde{\alpha}}}$, we can rewrite the system as:

$$\tilde{L}_{Ri} = \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j \tilde{L}_{Fj}^{\tilde{\alpha}} \tilde{L}_{Ri}^{\tilde{\beta}}$$

$$\tilde{L}_{Fj} = \sum_i d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j \tilde{L}_{Ri}^{\tilde{\beta}} \tilde{L}_{Fj}^{\tilde{\alpha}}$$

as desired.

□

The population and employment variables, $\tilde{L}_{Ri}$ and $\tilde{L}_{Fj}$, are now expressed as a product of the true population and employment variables, L_{Ri} and L_{Fj}, and the endogenous scalars κ_R and κ_F, which depend on the population-mobility assumptions. These equations are therefore useful for studying relative changes in population (e.g. L_{R1}/L_{R2}) and employment (e.g. L_{F1}/L_{F2}).

For reference, we state the theorem Allen et al. (2024), as applied to our model.

Allen et al. (2024) Remark 1 Suppose there exists a 2-by-2 matrix $\mathbf{A}$ such that for all $i \in \mathcal{S}$, and $y_{Ri} \in \mathbb{R}_{++}, y_{Fi} \in \mathbb{R}_{++}$,

$$\sum_k \left| \frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Fj}^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}}}}{\partial \ln y_{Rk}} \right| \leq (\mathbf{A})_{11}.$$

$$\sum_k \left| \frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Fj}^{\frac{\tilde{\alpha}}{1-\tilde{\alpha}}}}{\partial \ln y_{Fk}} \right| \leq (\mathbf{A})_{12}.$$

$$\sum_k \left| \frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Ri}^{\frac{\tilde{\beta}}{1-\tilde{\beta}}}}{\partial \ln y_{Rk}} \right| \leq (\mathbf{A})_{21}.$$

$$\sum_k \left| \frac{\partial \ln \sum_j d_{ij}^{-\phi} \tilde{U}_i \tilde{A}_j y_{Ri}^{\frac{\tilde{\beta}}{1-\tilde{\beta}}}}{\partial \ln y_{Fk}} \right| \leq (\mathbf{A})_{22}.$$

Then, if $\rho(\mathbf{A}) < 1$, there exists a unique solution to the system (49) and (50).

C General Equilibrium Effects for Residential Population

C.1 Proof of Theorem 5

Part (i) is an application of part 2 (ii.) in theorem 1.

For part (iii.), we first derive each term of the Neumann series matrix.

We are interested in the general equilibrium comparative statics expressed in equation (32) and repeated below:

$$\frac{\partial \ln \tilde{L}_{Rl}}{\partial \ln d_{ij}} = \frac{-L_{ij}\theta}{L_i^R} \underbrace{\left(\tilde{A}_{li}^0 + \tilde{A}_{li}^1 + ...\right)}_{A_{li}^{-1}} + \frac{-L_{ij}\theta}{L_j^F} \underbrace{\left(\tilde{A}_{l,N+j}^0 + \tilde{A}_{l,N+j}^1 + ...\right)}_{A_{l,N+j}^{-1}} \tag{51}$$

where $\tilde{A}_{li}^k$ is the (l,i) element of matrix $\tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}_X^k$.

We showed that we can express the inverse matrix $\mathbf{A}^{-1}$ as a series of matrices:

$$\mathbf{A}^{-1} = \tilde{\mathbf{\Gamma}}^{-1} \sum_{k=0}^{\infty} \boldsymbol{\alpha}_X^k$$
$$= \begin{bmatrix} \frac{1}{1-\beta\theta}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \frac{1}{1-\alpha\theta}\mathbf{I} \end{bmatrix} \sum_{k=0}^{\infty} \begin{bmatrix} \mathbf{0} & \frac{\alpha\theta}{1-\alpha\theta}\frac{\mathbf{L}}{\mathbf{L_R}} \\ \frac{\beta\theta}{1-\beta\theta}\frac{\mathbf{L^T}}{\mathbf{L_F}} & \mathbf{0} \end{bmatrix}^k$$

We seek to understand the remainder term

$$\mathbf{R}^1 \equiv \tilde{\mathbf{\Gamma}}^{-1} \sum_{k=1}^{\infty} \boldsymbol{\alpha}_X^k$$
$$= \begin{bmatrix} \frac{1}{1-\beta\theta}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \frac{1}{1-\alpha\theta}\mathbf{I} \end{bmatrix} \sum_{k=1}^{\infty} \begin{bmatrix} \mathbf{0} & \frac{\alpha\theta}{1-\alpha\theta}\frac{\mathbf{L}}{\mathbf{L_R}} \\ \frac{\beta\theta}{1-\beta\theta}\frac{\mathbf{L^T}}{\mathbf{L_F}} & \mathbf{0} \end{bmatrix}^k$$

Expansion of the Neumann series reveals that due to the block off-diagonal structure of matrix $\boldsymbol{\alpha}_X$, the even-degree terms have the structure:

$$\tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}_X^{2n} = \begin{bmatrix} \frac{1}{(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^n \left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^n & \mathbf{0} \\ \mathbf{0} & \frac{1}{(1-\alpha\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\right)^n \left(\frac{\mathbf{L^T}}{\mathbf{L_F}}\frac{\mathbf{L}}{\mathbf{L_R}}\right)^n \end{bmatrix}$$

while the odd terms have the structure:

$$\tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}_X^{2n+1} = \begin{bmatrix} \mathbf{0} & \frac{\alpha\theta}{(1-\beta\theta)(1-\alpha\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\right)^n \frac{\mathbf{L}}{\mathbf{L_R}}\left(\frac{\mathbf{L^T}}{\mathbf{L_F}}\frac{\mathbf{L}}{\mathbf{L_R}}\right)^n \\ \frac{\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^n \frac{\mathbf{L^T}}{\mathbf{L_F}}\left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^n & \mathbf{0} \end{bmatrix}$$

where $n = 0, 1, 2...$, and $2n$ are the even-degree terms and $2n+1$ are the odd-degree terms.

Combining the odd and even terms, we have:

$$\mathbf{A}^{-1} = \sum_{k=0}^{\infty} \tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}_X^k$$
$$= \sum_{n=0}^{\infty} \begin{bmatrix} \mathbf{0} & \frac{\alpha\theta}{(1-\beta\theta)(1-\alpha\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\right)^n \frac{\mathbf{L}}{\mathbf{L_R}}\left(\frac{\mathbf{L^T}}{\mathbf{L_F}}\frac{\mathbf{L}}{\mathbf{L_R}}\right)^n \\ \frac{\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^n \frac{\mathbf{L^T}}{\mathbf{L_F}}\left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^n & \mathbf{0} \end{bmatrix}$$
$$+ \begin{bmatrix} \frac{1}{(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^n \left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^n & \mathbf{0} \\ \mathbf{0} & \frac{1}{(1-\alpha\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\right)^n \left(\frac{\mathbf{L^T}}{\mathbf{L_F}}\frac{\mathbf{L}}{\mathbf{L_R}}\right)^n \end{bmatrix}$$
$$= \sum_{n=0}^{\infty} \begin{bmatrix} \frac{1}{(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^n \left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^n & \frac{\alpha\theta}{(1-\beta\theta)(1-\alpha\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\right)^n \frac{\mathbf{L}}{\mathbf{L_R}}\left(\frac{\mathbf{L^T}}{\mathbf{L_F}}\frac{\mathbf{L}}{\mathbf{L_R}}\right)^n \\ \frac{\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^n \frac{\mathbf{L^T}}{\mathbf{L_F}}\left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^n & \frac{1}{(1-\alpha\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\right)^n \left(\frac{\mathbf{L^T}}{\mathbf{L_F}}\frac{\mathbf{L}}{\mathbf{L_R}}\right)^n \end{bmatrix}$$

The term $\left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)$ captures the residential-to-firm-to-residential length-2 propogation. A shock to population affects employment locations proportionally to their probabilities of sourcing commuters, $\frac{\mathbf{L^T}}{\mathbf{L_F}}$, which then feeds back to residential locations proportionally to those locations probability to commuting to employment locations, $\frac{\mathbf{L}}{\mathbf{L_R}}$. The constant $\left(\frac{\alpha\theta\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\right)$ determines the extent of dampening of the shock over the 2-length propogation.

In our analysis we emphasize the distinction between the zeroth term and the remaining terms. The zeroth term is:

$$\tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}_X^0 = \begin{bmatrix} \frac{1}{1-\beta\theta}\mathbf{I} & \mathbf{0} \\ \\ \mathbf{0} & \frac{1}{1-\alpha\theta}\mathbf{I} \end{bmatrix}$$

Beyond the zeroth degree term, the series can be expressed as:

$$\sum_{k=1}^{\infty}\tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}_X^k = \sum_{n=0}^{\infty}\begin{bmatrix} \frac{1}{(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^{n+1}\left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^{n+1} & \frac{\alpha\theta}{(1-\beta\theta)(1-\alpha\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\right)^{n}\frac{\mathbf{L}}{\mathbf{L_R}}\left(\frac{\mathbf{L^T}}{\mathbf{L_F}}\frac{\mathbf{L}}{\mathbf{L_R}}\right)^{n} \\ \frac{\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^{n}\frac{\mathbf{L^T}}{\mathbf{L_F}}\left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^{n} & \frac{1}{(1-\alpha\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\right)^{n+1}\left(\frac{\mathbf{L^T}}{\mathbf{L_F}}\frac{\mathbf{L}}{\mathbf{L_R}}\right)^{n+1} \end{bmatrix}$$

For the residential impact, we focus on the top block, since only these elements appear in the residential comparative statics.

The top left block can be expressed as follows by incrementing the summation index.

$$\begin{aligned} R_{ij} &= \sum_{n=0}^{\infty}\frac{1}{(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^{n+1}\left[\left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^{n+1}\right]_{i,j} \\ &= \sum_{n=1}^{\infty}\frac{1}{(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^{n}\left[\left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^{n}\right]_{i,j} \end{aligned}$$

for $i \leq N, j \leq N$.

For the top right block, we have:

$$R^1_{ij} = \sum_{n=0}^{\infty}\frac{\alpha\theta}{(1-\beta\theta)(1-\alpha\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\alpha\theta)(1-\beta\theta)}\right)^{n}\left[\frac{\mathbf{L}}{\mathbf{L_R}}\left(\frac{\mathbf{L^T}}{\mathbf{L_F}}\frac{\mathbf{L}}{\mathbf{L_R}}\right)^{n}\right]_{i,j}$$

for $i \leq N, j > N$

C.2 Bounds on General Equilibrium Effects

We assume that $\alpha\beta > 0$ Consider first the top left block:

$$\sum_{n=0}^{\infty}\frac{1}{(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^{n+1}\left[\left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^{n+1}\right]_{i,j} = \sum_{n=1}^{\infty}\frac{1}{(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^{n}\left[\left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^{n}\right]_{i,j}$$
$$\leq \sum_{n=1}^{\infty}\frac{1}{(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^{n}$$
$$= \frac{1}{(1-\beta\theta)}\frac{\gamma^2}{1-\gamma^2}$$

where $\gamma^2 = \frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}$. The inequality in the second line holds because each term, $\left[\left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^{n}\right]_{i,j}$ is less than one, as each term is a weighted average of terms that are all less than one. The last line is a result of the geometric series, which converges when $\gamma^2 < 1$. We also have that $\sum_{n=0}^{\infty}\frac{1}{(1-\beta\theta)}\left(\frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\right)^{n+1}\left[\left(\frac{\mathbf{L}}{\mathbf{L_R}}\frac{\mathbf{L^T}}{\mathbf{L_F}}\right)^{n+1}\right]_{i,j} \geq 0$, since all terms are non-negative.

Similar expressions can be derived for the other blocks such that we have:

$$0 \leq R^1_{i,j} \leq \frac{1}{(1-\beta\theta)}\frac{\gamma^2}{1-\gamma^2} \quad \text{if } i \leq N, j \leq N$$
$$0 \leq R^1_{i,j} \leq \frac{1}{(1-\alpha\theta)}\frac{\gamma^2}{1-\gamma^2} \quad \text{if } i > N, j > N$$
$$\frac{\alpha\theta}{(1-\beta\theta)(1-\alpha\theta)}\frac{1}{1-\gamma^2} \leq R^1_{i,j} \leq 0 \quad \text{if } i \leq N, j > N, \alpha < 0$$
$$\frac{\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}\frac{1}{1-\gamma^2} \leq R^1_{i,j} \leq 0 \quad \text{if } i > N, j \leq N, \alpha < 0$$

C.3 Bounds on Symmetric Improvement

Without loss of generality, we now consider the relative population changes in locations 1 and 2 for a symmetric improvement, where $d\ln d_{12} = d\ln d_{21}$. Defining the series $R^K_{i,j} \equiv \left[\sum_{k=K}^{\infty}\tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}^k_X\right]_{ij}$ as the (i,j) element of the remaining Neumann series from K onwards. The first-order effect to the relative residential population of location 1 can be expressed as:

$$\frac{\partial\ln\tilde{L}_{R1}}{\partial\ln d_{12}} = \left[\frac{-L_{12}\theta}{L_{R1}}\underbrace{\left(\frac{1}{1-\beta\theta}+R^1_{11}\right)}_{A^{-1}_{11}} + \frac{-L_{12}\theta}{L_{F2}}\underbrace{\left(0+R^1_{1,N+2}\right)}_{A^{-1}_{1,N+2}}\right]$$

For a commuting cost in the reverse direction, $d\ln d_{21}$, the normalize increase in population is:

$$\frac{\partial \ln \tilde{L}_{R1}}{\partial \ln d_{21}} = \left[\frac{-L_{21}\theta}{L_{R2}} \underbrace{\left(0 + R^1_{12}\right)}_{A^{-1}_{12}} + \frac{-L_{21}\theta}{L_{F1}} \underbrace{\left(0 + R^1_{1,N+1}\right)}_{A^{-1}_{1,N+1}} \right]$$

Similarly, for location 2, the normalized comparative statics are given by:

$$\frac{\partial \ln \tilde{L}_{R2}}{\partial \ln d_{12}} = \left[\frac{-L_{12}\theta}{L_{R1}} \underbrace{\left(0 + R^1_{21}\right)}_{A^{-1}_{21}} + \frac{-L_{12}\theta}{L_{F2}} \underbrace{\left(0 + R^1_{2,N+2}\right)}_{A^{-1}_{2,N+2}} \right]$$

$$\frac{\partial \ln \tilde{L}_{R2}}{\partial \ln d_{21}} = \left[\frac{-L_{21}\theta}{L_{R2}} \underbrace{\left(\frac{1}{1-\beta\theta} + R^1_{22}\right)}_{A^{-1}_{22}} + \frac{-L_{21}\theta}{L_{F1}} \underbrace{\left(0 + R^1_{2,N+1}\right)}_{A^{-1}_{2,N+1}} \right]$$

For a symmetric transport improvement, where $d \ln d_{12} = d \ln d_{21}$ the relative population changes are therefore:

$$\begin{aligned} \frac{\partial \ln(L_{R1}/L_{R2})}{\partial \ln d_{12}} + \frac{\partial \ln(L_{R1}/L_{R2})}{\partial \ln d_{21}} &= \frac{-\theta}{1-\beta\theta} \left(\frac{L_{12}}{L_{R1}} - \frac{L_{21}}{L_{R2}} \right) \\ &+ \frac{-L_{12}\theta}{L_{R1}} \left(R^1_{1,1} - R^1_{2,1} \right) \\ &+ \frac{-L_{12}\theta}{L_{F2}} \left(R^1_{1,N+2} - R^1_{2,N+2} \right) \\ &+ \frac{-L_{21}\theta}{L_{R2}} \left(R^1_{1,2} - R^1_{2,2} \right) \\ &+ \frac{-L_{21}\theta}{L_{F1}} \left(R^1_{1,N+1} - R^1_{2,N+1} \right) \end{aligned} \tag{52}$$

For each term in (52) and using the above bounds, we have:

$$\begin{aligned}
\left|\frac{L_{12}\theta}{L_{R1}}\left(R^1_{1,1}-R^1_{2,1}\right)\right| &\leq \frac{L_{12}\theta}{L_{R1}}\frac{1}{(1-\beta\theta)}\frac{\gamma^2}{1-\gamma^2}\\
\left|\frac{L_{21}\theta}{L_{R2}}\left(R^1_{1,2}-R^1_{2,2}\right)\right| &\leq \frac{L_{21}\theta}{L_{R2}}\frac{1}{(1-\beta\theta)}\frac{\gamma^2}{1-\gamma^2}\\
\left|\frac{L_{12}\theta}{L_{F2}}\left(R^1_{1,N+2}-R^1_{2,N+2}\right)\right| &\leq \frac{L_{12}\theta}{L_{F2}}\frac{|\alpha\theta|}{(1-\beta\theta)(1-\alpha\theta)}\frac{1}{1-\gamma^2}\\
\left|\frac{L_{21}\theta}{L_{F1}}\left(R^1_{1,N+1}-R^1_{2,N+1}\right)\right| &\leq \frac{L_{21}\theta}{L_{F1}}\frac{|\alpha\theta|}{(1-\beta\theta)(1-\alpha\theta)}\frac{1}{1-\gamma^2}
\end{aligned}$$

We seek sufficient conditions for the symmetric improvement to lead to a relative increase in population in location 1 relative to location 2, i.e. $-\left[\frac{\partial\ln(L_{R1}/L_{R2})}{\partial\ln d_{12}}+\frac{\partial\ln(L_{R1}/L_{R2})}{\partial\ln d_{21}}\right]\geq 0$. Thus:

$$\begin{aligned}
-\left[\frac{\partial\ln(L_{R1}/L_{R2})}{\partial\ln d_{12}}+\frac{\partial\ln(L_{R1}/L_{R2})}{\partial\ln d_{21}}\right] &= \frac{\theta}{1-\beta\theta}\left(\frac{L_{12}}{L_{R1}}-\frac{L_{21}}{L_{R2}}\right)\\
&\quad+\frac{L_{12}\theta}{L_{R1}}\left(R^1_{1,1}-R^1_{2,1}\right)\\
&\quad+\frac{L_{12}\theta}{L_{F2}}\left(R^1_{1,N+2}-R^1_{2,N+2}\right)\\
&\quad+\frac{L_{21}\theta}{L_{R2}}\left(R^1_{1,2}-R^1_{2,2}\right)\\
&\quad+\frac{L_{21}\theta}{L_{F1}}\left(R^1_{1,N+1}-R^1_{2,N+1}\right)\\
&\geq 0
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow \frac{\theta}{1-\beta\theta}\left(\frac{L_{12}}{L_{R1}}-\frac{L_{21}}{L_{R2}}\right) &\geq \frac{L_{12}\theta}{L_{R1}}\left(R^1_{1,1}-R^1_{2,1}\right)\\
&\quad+\frac{L_{12}\theta}{L_{F2}}\left(R^1_{1,N+2}-R^1_{2,N+2}\right)\\
&\quad+\frac{L_{21}\theta}{L_{R2}}\left(R^1_{1,2}-R^1_{2,2}\right)\\
&\quad+\frac{L_{21}\theta}{L_{F1}}\left(R^1_{1,N+1}-R^1_{2,N+1}\right)
\end{aligned}$$

Then, since

$$
\begin{aligned}
&\left|\frac{L_{12}\theta}{L_{R1}}\left(R^1_{1,1}-R^1_{2,1}\right)\right| \\
&+\left|\frac{L_{12}\theta}{L_{F2}}\left(R^1_{1,N+2}-R^1_{2,N+2}\right)\right| \\
&+\left|\frac{L_{21}\theta}{L_{R2}}\left(R^1_{1,2}-R^1_{2,2}\right)\right| \\
&+\left|\frac{L_{21}\theta}{L_{F1}}\left(R^1_{1,N+1}-R^1_{2,N+1}\right)\right| \\
&\geq \frac{L_{12}\theta}{L_{R1}}\left(R^1_{1,1}-R^1_{2,1}\right) \\
&+\frac{L_{12}\theta}{L_{F2}}\left(R^1_{1,N+2}-R^1_{2,N+2}\right) \\
&+\frac{L_{21}\theta}{L_{R2}}\left(R^1_{1,2}-R^1_{2,2}\right) \\
&+\frac{L_{21}\theta}{L_{F1}}\left(R^1_{1,N+1}-R^1_{2,N+1}\right)
\end{aligned}
$$

a sufficient condition for a population shift is::

$$
\begin{aligned}
\frac{\theta}{1-\beta\theta}\left(\frac{L_{12}}{L_{R1}}-\frac{L_{21}}{L_{R2}}\right) \geq{}& \frac{L_{12}\theta}{L_{R1}}\frac{1}{(1-\beta\theta)}\frac{\gamma^2}{1-\gamma^2} \\
&+\frac{L_{21}\theta}{L_{R2}}\frac{1}{(1-\beta\theta)}\frac{\gamma^2}{1-\gamma^2} \\
&+\frac{L_{12}\theta}{L_{F2}}\frac{|\alpha\theta|}{(1-\beta\theta)(1-\alpha\theta)}\frac{1}{1-\gamma^2} \\
&+\frac{L_{21}\theta}{L_{F1}}\frac{|\alpha\theta|}{(1-\beta\theta)(1-\alpha\theta)}\frac{1}{1-\gamma^2}
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow \frac{L_{12}}{L_{R1}}\left(1-\frac{\gamma^2}{1-\gamma^2}\right) \geq{}& \frac{L_{21}}{L_{R2}}\left(1+\frac{\gamma^2}{1-\gamma^2}\right) \\
&+\frac{L_{12}}{L_{F2}}\frac{|\alpha|}{(1-\alpha)}\frac{1}{1-\gamma^2} \\
&+\frac{L_{21}}{L_{F1}}\frac{|\alpha|}{(1-\alpha)}\frac{1}{1-\gamma^2}
\end{aligned}
$$

$$\Leftrightarrow \frac{L_{12}}{L_{R1}}\left(\frac{1-2\gamma^2}{1-\gamma^2}\right) \geq \frac{L_{21}}{L_{R2}}\frac{1}{1-\gamma^2} + \frac{L_{12}}{L_{F2}}\frac{|\alpha|}{(1-\alpha)}\frac{1}{1-\gamma^2} + \frac{L_{21}}{L_{F1}}\frac{|\alpha|}{(1-\alpha)}\frac{1}{1-\gamma^2}$$

$$\Leftrightarrow \frac{L_{12}}{L_{R1}} \geq \frac{L_{21}}{L_{R2}}\frac{1}{1-2\gamma^2} + \left(\frac{L_{12}}{L_{F2}} + \frac{L_{21}}{L_{F1}}\right)\frac{|\alpha|}{(1-\alpha)}\frac{1}{1-2\gamma^2}$$

$$\Leftrightarrow \frac{L_{12}}{L_{R1}} \geq \frac{1}{1-2\gamma^2}\left[\frac{L_{21}}{L_{R2}} + \left(\frac{L_{12}}{L_{F2}} + \frac{L_{21}}{L_{F1}}\right)\frac{|\alpha|}{(1-\alpha)}\right]$$

One can rearrange the above expression to obtain $\frac{L_{ij}}{L_{ji}} > \frac{L_{Ri}L_{Fj}(L_{Fi}+L_{Rj}\bar{\alpha})}{L_{Rj}L_{Fi}(L_{Fj}(1-2\gamma^2)-L_{Ri}\bar{\alpha})}$

C.4 Closed-Form Expression for Remainders for Symmetric improvement

In this section, we show that Given that $\mathbf{A}^{-1} = \sum_{k=0}^{\infty} \tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}_X^k$ and the remainder is $\mathbf{R}^1 = \sum_{k=1}^{\infty} \tilde{\mathbf{\Gamma}}^{-1}\boldsymbol{\alpha}_X^k$, $\boldsymbol{\alpha}_X$ can be factored out to express the remainder as $\mathbf{R}^1 = \tilde{\mathbf{\Gamma}}^{-1}\sum_{k=1}^{\infty}\boldsymbol{\alpha}_X^k = \tilde{\mathbf{\Gamma}}^{-1}\left[\sum_{k=0}^{\infty}\boldsymbol{\alpha}_X^k\right]\boldsymbol{\alpha}_X = \mathbf{A}^{-1}\boldsymbol{\alpha}_X$, which can be multiplied by $\mathbf{T}$ expressed in (26) to obtain the 1st-degree remainder for all comparative statics, or more explicitly we can substitute into (52).

Another approach is to make each general equilibrium force explicit. Defining $\mathbf{M} \equiv \frac{\mathbf{L}}{\mathbf{L_R}} \cdot \frac{\mathbf{L}^T}{\mathbf{L_F}}$ and $c \equiv \frac{\alpha\theta\beta\theta}{(1-\beta\theta)(1-\alpha\theta)}$, we have the following relationships for the remainder terms:

$$R^1_{1,1} - R^1_{2,1} = \frac{1}{1-\beta\theta}\sum_{n=1}^{\infty} c^n \left([\mathbf{M}^n]_{11} - [\mathbf{M}^n]_{21}\right)$$

$$R^1_{1,2} - R^1_{2,2} = \frac{1}{1-\beta\theta}\sum_{n=1}^{\infty} c^n \left([\mathbf{M}^n]_{12} - [\mathbf{M}^n]_{22}\right)$$

$$R^1_{1,N+2} - R^1_{2,N+2} = \frac{\alpha\theta}{(1-\beta\theta)(1-\alpha\theta)}\sum_{n=1}^{\infty} c^n \left(\left[\frac{\mathbf{L}}{\mathbf{L_R}}\mathbf{M}^n\right]_{1,N+2} - \left[\frac{\mathbf{L}}{\mathbf{L_R}}\mathbf{M}^n\right]_{2,N+2}\right)$$

$$R^1_{1,N+1} - R^1_{2,N+1} = \frac{\alpha\theta}{(1-\beta\theta)(1-\alpha\theta)} \sum_{n=1}^{\infty} c^n \left(\left[\frac{\mathbf{L}}{\mathbf{L_R}} \mathbf{M}^n \right]_{1,N+1} - \left[\frac{\mathbf{L}}{\mathbf{L_R}} \mathbf{M}^n \right]_{2,N+1} \right)$$

If the spectral radius $\rho(cM) < 1$, then:

$$\sum_{n=1}^{\infty} c^n \mathbf{M}^n = c\mathbf{M}(\mathbf{I} - c\mathbf{M})^{-1}$$

Then defining v_{ij} as the row vector with the i^{th} element equal to 1, the j^{th} element equal to -1, and all other elements equal to 0, and e_i as the column vector with the i^{th} element equal to 1 and all other elements equal to 0, we have:

$$R^1_{1,1} - R^1_{2,1} = \frac{c}{1-\beta\theta} \cdot v_{12} \mathbf{M}(\mathbf{I} - c\mathbf{M})^{-1} e_1$$

$$R^1_{1,2} - R^1_{2,2} = \frac{c}{1-\beta\theta} \cdot v_{12} \mathbf{M}(\mathbf{I} - c\mathbf{M})^{-1} e_2$$

$$R^1_{1,N+2} - R^1_{2,N+2} = \frac{\alpha\theta c}{(1-\beta\theta)(1-\alpha\theta)} \cdot v_{12} \frac{\mathbf{L}}{\mathbf{L_R}} \mathbf{M}(\mathbf{I} - c\mathbf{M})^{-1} e_{N+2}$$

$$R^1_{1,N+1} - R^1_{2,N+1} = \frac{\alpha\theta c}{(1-\beta\theta)(1-\alpha\theta)} \cdot v_{12} \frac{\mathbf{L}}{\mathbf{L_R}} \mathbf{M}(\mathbf{I} - c\mathbf{M})^{-1} e_{N+1}$$

D Comparative statics for Ahlfeldt et al. (2015)

We derive first-order comparative statics expressions for the model similar to Ahlfeldt et al. (2015). We consider the case where $\bar{A}_i > 0, \bar{U}_i > 0$ so that equilibrium variables are strictly positive. We first consider a case where floorspace is exogenously separated between residential and commercial uses. Let H_{Ri} denote the exogenous amount of residential floorspace and let H_{Fi} denote the exonenous amount of commercial floorspace. Consider the system of equations:

$$L_{Ri} = \bar{L}\left(\frac{\gamma}{\bar{V}}\right)^{\theta}\left(U_i\, r_i^{\beta-1}\right)^{\theta}\sum_j\left(\frac{w_j}{d_{ij}}\right)^{\theta}, \tag{LR}$$

$$L_{Fj} = \bar{L}\left(\frac{\gamma}{\bar{V}}\right)^{\theta} w_j^{\theta}\sum_i\left(\frac{U_i}{r_i^{1-\beta} d_{ij}}\right)^{\theta}, \tag{LF}$$

$$\bar{y}_i = \sum_j \frac{w_j (w_j/d_{ij})^{\theta}}{\Phi_{Ri}}, \qquad \Phi_{Ri} = \sum_j (w_j/d_{ij})^{\theta}, \tag{$\bar{y}$}$$

$$\bar{V} = \gamma\left[\sum_j\sum_i\sum_m\left(\frac{U_i\, w_j\, r_i^{\beta-1}}{d_{ijm}}\right)^{\theta}\right]^{1/\theta}, \tag{$\bar{V}$}$$

$$H_i \upsilon_{Ri} = \frac{(1-\beta)\,\bar{y}_i\, L_{Ri}}{r_i}, \qquad H_i\upsilon_{Fi} = L_{Fi}\left(\frac{A_i(1-\alpha)}{r_i}\right)^{1/\alpha}, \qquad \upsilon_{Ri}+\upsilon_{Fi} = 1, \tag{Land}$$

$$L_{Fi} = H_i \upsilon_{Fi}\left(\frac{A_i \alpha}{w_i}\right)^{1/(1-\alpha)}. \tag{FD}$$

$$U_i = \bar{U}_i\, L_{Ri}^{\mu_U}, \qquad A_i = \bar{A}_i\, L_{Fi}^{\mu_A} \tag{ext}$$

Next, we consider the case where floorspace is efficiently shared between commercial and residential uses and floorprices are equalized. Before proceeding, we first show that the model is homogenous in the total population $\bar{L}$.

Lemma 1 (Homogeneity in $\bar{L}$ with fixed H_i and land shares)**.** Consider $\theta > 0$, $\alpha \in (0,1)$, $\beta \in (0,1)$, and the system:

$$L_{Ri} = \bar{L}\left(\frac{\gamma}{\bar{V}}\right)^{\theta}\left(U_i\, r_i^{\beta-1}\right)^{\theta}\sum_j\left(\frac{w_j}{d_{ij}}\right)^{\theta}, \tag{LR}$$

$$L_{Fj} = \bar{L}\left(\frac{\gamma}{\bar{V}}\right)^{\theta} w_j^{\theta}\sum_i\left(\frac{U_i}{r_i^{1-\beta} d_{ij}}\right)^{\theta}, \tag{LF}$$

$$\bar{y}_i = \sum_j \frac{w_j (w_j/d_{ij})^{\theta}}{\Phi_{Ri}}, \qquad \Phi_{Ri} = \sum_j (w_j/d_{ij})^{\theta}, \tag{$\bar{y}$}$$

$$\bar{V} = \gamma\left[\sum_j\sum_i\sum_m\left(\frac{U_i\, w_j\, r_i^{\beta-1}}{d_{ijm}}\right)^{\theta}\right]^{1/\theta}, \tag{$\bar{V}$}$$

$$H_i \upsilon_{Ri} = \frac{(1-\beta)\,\bar{y}_i\, L_{Ri}}{r_i}, \qquad H_i\upsilon_{Fi} = L_{Fi}\left(\frac{A_i(1-\alpha)}{r_i}\right)^{1/\alpha}, \qquad \upsilon_{Ri}+\upsilon_{Fi} = 1, \tag{Land}$$

$$L_{Fi} = H_i \upsilon_{Fi}\left(\frac{A_i \alpha}{w_i}\right)^{1/(1-\alpha)}. \tag{FD}$$

$$U_i = \bar{U}_i\, L_{Ri}^{\mu_U}, \qquad A_i = \bar{A}_i\, L_{Fi}^{\mu_A} \tag{ext}$$

Suppose $H_i > 0$ are fixed and $(L_{Ri}, L_{Fj}, w_j, r_i, \bar{V}, \upsilon_{Ri}, \upsilon_{Fi})$ solves (LR)–(ext) at some $\bar{L} > 0$.

For any $s > 0$, define the scaled variables

$$\begin{aligned}
&\bar{L}' = s\,\bar{L}, \quad L'_{Ri} = s\,L_{Ri}, \quad L'_{Fj} = s\,L_{Fj}, \\
&r'_i = s^{\,b} r_i, \quad w'_j = s^{\,a} w_j, \quad \bar{V}' = s^{\,\nu}\bar{V}, \\
&U'_i = \bar{U}_i (L'_{Ri})^{\mu_U} = s^{\mu_U} U_i, \qquad A'_i = \bar{A}_i (L'_{Fi})^{\mu_A} = s^{\mu_A} A_i, \\
&\upsilon'_{Ri} = \upsilon_{Ri}, \quad \upsilon'_{Fi} = \upsilon_{Fi} \quad \text{(land shares unchanged).}
\end{aligned} \tag{53}$$

with exponents

$$b = \mu_A + \alpha, \qquad a = \mu_A + \alpha - 1, \qquad \nu = \mu_U + (\mu_A + \alpha)\beta - 1. \tag{54}$$

Then $(L'_{Ri}, L'_{Fj}, w'_j, r'_i, \bar{V}', \upsilon'_{Ri}, \upsilon'_{Fi})$ together with the same fixed $\{H_i\}$ satisfies (LR)–(ext) at $\bar{L}'$, and the spatial shares are invariant:

$$\frac{L'_{Ri}}{\bar{L}'} = \frac{L_{Ri}}{\bar{L}}, \qquad \frac{L'_{Fj}}{\bar{L}'} = \frac{L_{Fj}}{\bar{L}}, \qquad (\upsilon'_{Ri}, \upsilon'_{Fi}) = (\upsilon_{Ri}, \upsilon_{Fi}).$$

Hence the equilibrium population and employment is homogeneous of degree one in $\bar{L}$ (with fixed H_i).

Proof. All statements follow by direct scaling.

(A) Floor-space identities with fixed H_i. Using (53), $\bar{y}'_i = s^a \bar{y}_i$ (from $(\bar{y})$ being homogeneous of degree a in w), and $r'_i = s^b r_i$, $L'_{Ri} = sL_{Ri}$:

$$\frac{(1-\beta)\bar{y}'_i L'_{Ri}}{r'_i} = s^{a+1-b} \frac{(1-\beta)\bar{y}_i L_{Ri}}{r_i}.$$

Likewise for the commercial term in (Land):

$$L'_{Fi}\left(\frac{A'_i(1-\alpha)}{r'_i}\right)^{1/\alpha} = s^{\,1+\frac{\mu_A - b}{\alpha}} L_{Fi}\left(\frac{A_i(1-\alpha)}{r_i}\right)^{1/\alpha}.$$

With the choices (54), both exponents vanish because $a + 1 - b = 0$ and $1 + (\mu_A - b)/\alpha = 0$. Hence each side of (Land) is unchanged, so $H_i \upsilon_{Ri}$ and $H_i \upsilon_{Fi}$ are fixed and, with H_i fixed, $(\upsilon_{Ri}, \upsilon_{Fi})$ are unchanged.

(B) Firm labor demand (FD). The right-hand side scales as $s^{(\mu_A - a)/(1-\alpha)}$. Requiring $L'_{Fi} = sL_{Fi}$ forces

$$1 = \frac{\mu_A - a}{1-\alpha} \quad \Longleftrightarrow \quad a = \mu_A + \alpha - 1,$$

which matches (54).

(C) Location equations (LR)–(LF)**.** In (LR), the right-hand side scales by

$$s^{\,1\,-\nu\theta\,+\,\theta\mu_U\,+\,\theta b(\beta-1)\,+\,a\theta}.$$

Requiring $L'_{Ri} = sL_{Ri}$ gives

$$-\nu + \mu_U + b(\beta-1) + a = 0. \qquad (R)$$

In (LF), the right-hand side scales by

$$s^{\,1\,-\nu\theta\,+\,a\theta\,+\,\theta\mu_U\,-\,\theta b(1-\beta)},$$

and $L'_{Fj} = sL_{Fj}$ gives

$$-\nu + a + \mu_U - b(1-\beta) = 0. \qquad (F)$$

With (54), both (R) and (F) hold identically.

(D) Price index $(\bar{V})$**.** Inside the bracket, each term scales as $s^{\theta(\mu_U+a+b(\beta-1))}$, so $\bar{V}' = \gamma[\cdots]^{1/\theta}$ scales as $s^{\mu_U+a+b(\beta-1)}\bar{V}$. This equals $s^{\nu}\bar{V}$ by (54), so $(\bar{V})$ is satisfied.

Combining (A)–(D), all equilibrium conditions hold at $\bar{L}'$, with invariant spatial shares and land shares. □

Corollary 1 (Homogeneity with rent–elastic land supply)**.** Augment the system in Lemma 1 by replacing the fixed-H_i assumption with the supply schedule $H_i = \bar{H}_i\, r_i^{\varepsilon}$, $\varepsilon \geq 0$.

For any $s > 0$, consider the change of variables

$$\bar{L} \mapsto s\,\bar{L}, \quad L_{Ri} \mapsto s\,L_{Ri}, \quad L_{Fi} \mapsto s\,L_{Fi}, \quad r_i \mapsto s^{\,b}r_i, \quad w_i \mapsto s^{\,a}w_i, \quad \bar{V} \mapsto s^{\,\nu}\bar{V}, \quad U_i \mapsto s^{\mu_U}U_i, \quad A_i \mapsto s^{\mu_A}$$

with $\upsilon_{Ri}, \upsilon_{Fi}$ unchanged. Then the system remains satisfied with the spatial distribution provided the exponents are

$$b \;=\; \frac{\alpha+\mu_A}{1+\alpha\varepsilon}, \qquad a \;=\; \frac{\alpha+\mu_A(1+\varepsilon)-1}{1+\alpha\varepsilon}, \qquad \nu \;=\; \mu_U \;+\; \frac{\beta\alpha+\beta\mu_A+\mu_A\varepsilon-1}{1+\alpha\varepsilon}\;.$$

In particular, $\varepsilon = 0$ recovers Lemma 1: $b = \mu_A+\alpha$, $a = \mu_A+\alpha-1$, $\nu = \mu_U+(\mu_A+\alpha)\beta-1$. If productivity is exogenous ($\mu_A = 0$), these simplify to

$$b = \frac{\alpha}{1+\alpha\varepsilon}, \qquad a = \frac{\alpha-1}{1+\alpha\varepsilon}, \qquad \nu = \mu_U + \frac{\beta\alpha-1}{1+\alpha\varepsilon}.$$

Proof (check scaling). Write market clearing as $H_i = \underbrace{\frac{(1-\beta)\bar{y}_i L_{Ri}}{r_i}}_{H_{Ri}} + \underbrace{L_{Fi}\Big(\frac{A_i(1-\alpha)}{r_i}\Big)^{1/\alpha}}_{H_{Fi}}$ and $H_i = \bar{H}_i r_i^{\varepsilon}$. Under the map above:

$$H_i^S:\ s^{b\varepsilon}, \qquad H_{Ri}:\ s^{\,a+1-b}, \qquad H_{Fi}:\ s^{\,1+\frac{\mu_A - b}{\alpha}}.$$

To keep the land-use shares $\upsilon_{Ri}, \upsilon_{Fi}$ invariant, require each demand addend to scale like supply: $a+1-b = b\varepsilon$ and $1 + \frac{\mu_A - b}{\alpha} = b\varepsilon$. Solving gives

$$b = \frac{\alpha + \mu_A}{1+\alpha\varepsilon}, \qquad a = b(1+\varepsilon) - 1 = \frac{\alpha + \mu_A(1+\varepsilon) - 1}{1+\alpha\varepsilon}.$$

The firm labor demand $L_{Fi} = H_i \upsilon_{Fi}(A_i \alpha / w_i)^{1/(1-\alpha)}$ scales as $s^{\,b\varepsilon + \frac{\mu_A - a}{1-\alpha}}$ on the right; matching $L_{Fi} \mapsto sL_{Fi}$ yields $b\varepsilon + \frac{\mu_A - a}{1-\alpha} = 1$, which is satisfied by the a, b above. Finally, the resident/firm location equations imply $-\nu + \mu_U + b(\beta - 1) + a = 0$ (equivalently $-\nu + \mu_U + a - b(1-\beta) = 0$), giving $\nu = \mu_U + a + b(\beta - 1)$, i.e. the stated expression. The price-index equation then holds by homogeneity (as in the lemma). Hence shares are preserved. □

Corollary 2 (Normalization via change of variables)**.** Let the exponents (a, b, ν) be as in Lemma 1, i.e. $b = \mu_A + \alpha$, $a = \mu_A + \alpha - 1$, $\nu = \mu_U + (\mu_A + \alpha)\beta - 1$. Fix any constant $\kappa > 0$ and define

$$s \;\equiv\; \left(\frac{\kappa\, \bar{V}^{\theta}}{\bar{L}}\right)^{\frac{1}{1-\nu\theta}}.$$

Apply the one-dimensional change of variables

$$\tilde{L} = s\,\bar{L}, \quad \tilde{V} = s^{\nu}\bar{V}, \quad \tilde{w} = s^{a} w, \quad \tilde{r} = s^{b} r, \quad \tilde{L}_R = sL_R, \quad \tilde{L}_F = sL_F, \quad \tilde{U} = s^{\mu_U} U, \quad \tilde{A} = s^{\mu_A} A,$$

with $\tilde{\upsilon}_{Ri} = \upsilon_{Ri}$, $\tilde{\upsilon}_{Fi} = \upsilon_{Fi}$, and H_i fixed. Then

$$\frac{\tilde{L}}{\tilde{V}^{\theta}} = \kappa$$

Without loss of generality, setting $\kappa = 1$, we have a new system of equations:

$$\tilde{L}_{Ri} = \left(\tilde{U}_i \, \tilde{r}_i^{\beta-1}\right)^{\theta} \sum_j \left(\frac{\tilde{w}_j}{d_{ij}}\right)^{\theta} \tag{55}$$

$$\tilde{L}_{Fj} = \tilde{w}_j^{\theta} \sum_i \left(\frac{\tilde{U}_i}{\tilde{r}_i^{1-\beta} \, d_{ij}}\right)^{\theta} \tag{56}$$

$$\tilde{\bar{y}}_i = \sum_j \frac{\tilde{w}_j \left(\frac{\tilde{w}_j}{d_{ij}}\right)^{\theta}}{\tilde{\Phi}_{Ri}}, \qquad \tilde{\Phi}_{Ri} = \sum_j \left(\frac{\tilde{w}_j}{d_{ij}}\right)^{\theta} \tag{57}$$

$$H_i \, \tilde{\upsilon}_{Ri} = \frac{(1-\beta) \, \tilde{\bar{y}}_i \, \tilde{L}_{Ri}}{\tilde{r}_i}, \qquad H_i \, \tilde{\upsilon}_{Fi} = \tilde{L}_{Fi} \left(\frac{A_i(1-\alpha)}{\tilde{r}_i}\right)^{1/\alpha}, \qquad \tilde{\upsilon}_{Ri} + \tilde{\upsilon}_{Fi} = 1 \tag{58}$$

$$\tilde{L}_{Fi} = H_i \, \tilde{\upsilon}_{Fi} \left(\frac{A_i \, \alpha}{\tilde{w}_i}\right)^{1/(1-\alpha)} \tag{59}$$

$$\tilde{U}_i = \bar{U}_i \, (\tilde{L}_{Ri})^{\mu_U}, \qquad \tilde{A}_i = \bar{A}_i \, (\tilde{L}_{Fi})^{\mu_A} \tag{60}$$

where we no longer have an equation expressing the average utility.

An analogous argument can be made if there is a constant elasticity of supply of floorspace (from corollary 1)

D.1 Comparative Statics

We rearrange the equations above into the following system of equations:

$$\tilde{L}_{Ri}\left(\tilde{L}_{Ri}^{\mu_U}\,\tilde{r}_i^{\beta-1}\right)^{-\theta} = \bar{U}_i^{-\theta}\sum_j\left(\frac{\tilde{w}_j}{d_{ij}}\right)^{\theta} \tag{L_R}$$

$$\tilde{L}_{Fj}\tilde{w}_j^{-\theta} = \sum_i\left(\frac{\bar{U}_i\tilde{L}_{Ri}^{\mu_U}}{\tilde{r}_i^{1-\beta}\,d_{ij}}\right)^{\theta} \tag{L_F}$$

$$\tilde{\bar{y}}_i\tilde{L}_{Ri}\left(\tilde{L}_{Ri}^{\mu_U}\,\tilde{r}_i^{\beta-1}\right)^{-\theta} = \bar{U}_i^{-\theta}\sum_j\tilde{w}_j\left(\frac{\tilde{w}_j}{d_{ij}}\right)^{\theta}, \tag{$\bar{y}$}$$

$$\tilde{r}_i = \frac{(1-\beta)\,\tilde{\bar{y}}_i\,\tilde{L}_{Ri}}{\upsilon_{Ri}H_i}, \tag{r}$$

$$\upsilon_{Fi} = \frac{\tilde{L}_{Fi}}{H_i}\left(\frac{\bar{A}_i\tilde{L}_{Fi}^{\mu_A}(1-\alpha)}{\tilde{r}_i}\right)^{1/\alpha} \tag{υ_F}$$

$$\tilde{w}_i = \left(\frac{H_i\,\upsilon_{Fi}}{\tilde{L}_{Fi}}\right)^{1-\alpha}\bar{A}_iL_{Fi}^{\mu_A}\,\alpha \tag{w}$$

$$\upsilon_{Ri} = 1-\upsilon_{Fi} \tag{υ_{Ri}}$$

The first three equations are written in the constant elasticity form of equations (1). The next three equations are written in the constant elasticity form

$$x_{ih'} = K_{ih}\prod_{h'\in\mathcal{H}}x_{ih'}^{\sigma_{hh'}} \tag{61}$$

And the final equation is not in a constant-elasticity form but the endogenous variables are separable from any location-specific fundamentals that may be of interest for comparative statics (and in particular for our interest, d_{ij})

Let $\mathbf{x}_i = [\tilde{L}_{Ri}, \tilde{L}_{Fi}, \tilde{\bar{y}}_i, r_i, \upsilon_{Fi}, \tilde{w}_i, \upsilon_{Ri}]'$ denote our endogenous variables of interest and define the change of variables $\ln\mathbf{y}_i = \gamma\ln\mathbf{x}_i$ where

$$\boldsymbol{\gamma} = \begin{bmatrix} 1-\mu_U\theta & 0 & 0 & -(\beta-1)\theta & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -\theta & 0 \\ 1-\mu_U\theta & 0 & 1 & -(\beta-1)\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

so that the left-hand-side of our equations are our newly defined variables.

Let $\mathbf{B}$ denote the constant-elasticity matrix for the first three equations:

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \theta & 0 \\ 0 & 0 & 0 & \theta(\beta-1) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1+\theta & 0 \end{bmatrix}$$

$$\boldsymbol{\sigma} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 1+\frac{\mu_A}{\alpha} & 0 & -\frac{1}{\alpha} & 0 & 0 & 0 \\ 0 & \mu_A-(1-\alpha) & 0 & 0 & 1-\alpha & 0 & 0 \end{bmatrix}$$

so that we then have the system:

$$y_{ih} = \sum_{j\in\mathcal{N}} K_{ijh} \prod_{h'\in\mathcal{H}} y_{jh'}^{\alpha_{hh'}}, \qquad h = 1,2,3 \tag{62}$$

$$y_{ih} = K_{ih} \prod_{h'\in\mathcal{H}} y_{ih'}^{\beta_{hh'}}, \quad h = 4,5,6 \tag{63}$$

$$y_{i7} = 1 - y_{i5} \tag{64}$$

where $\alpha_{hh'}$ refers to the h,h' elements of the matrix $\boldsymbol{\alpha} = \mathbf{B}\boldsymbol{\gamma}^{-1}$ and $\beta_{hh'}$ refers to the h,h' elements of the matrix $\boldsymbol{\beta} = \boldsymbol{\sigma}\boldsymbol{\gamma}^{-1}$. We define the following:

$$\boldsymbol{\alpha}_X \equiv \text{bdiag}\Big(\mathbf{X}_{(1)}, \mathbf{X}_{(2)}, \mathbf{X}_{(3)}\Big)\Big(\boldsymbol{\alpha} \otimes \mathbf{I}_N\Big)$$

$$\Big[\mathbf{X}_{(1)}\Big]_{ij} = \frac{\left(\frac{\tilde{w}_j}{d_{ij}}\right)^{\theta}}{\sum_k \left(\frac{\tilde{w}_k}{d_{ik}}\right)^{\theta}}, \quad \Big[\mathbf{X}_{(2)}\Big]_{ji} = \frac{\left(\frac{\bar{U}_i\,\tilde{r}_i^{\beta-1}}{d_{ij}}\right)^{\theta}}{\sum_k \left(\frac{\bar{U}_k\,\tilde{r}_k^{\beta-1}}{d_{kj}}\right)^{\theta}}, \quad \Big[\mathbf{X}_{(3)}\Big]_{ij} = \frac{\tilde{w}_j^{1+\theta}\,d_{ij}^{-\theta}}{\sum_k \tilde{w}_k^{1+\theta}\,d_{ik}^{-\theta}}$$

$$\tilde{\boldsymbol{\upsilon}} \equiv \Big[\mathbf{0} \quad \mathbf{0} \quad \mathbf{0} \quad \mathbf{0} \quad \text{diag}\Big(\tfrac{-\boldsymbol{\upsilon}_F}{\boldsymbol{\upsilon}_R}\Big) \quad \mathbf{0} \quad \mathbf{0}\Big]$$

where we note that $\mathbf{X}_{(1)} = \dfrac{\mathbf{L}}{\mathbf{L_R}}$ is the residential commuter flow shares, $\mathbf{X}_{(2)} = \dfrac{\mathbf{L}^{\top}}{\mathbf{L_F}}$ is the firm commuter flow shares, and $\mathbf{X}_{(3)} = \dfrac{\mathbf{L}\circ\mathbf{w}'}{\mathbf{L_R}\circ\bar{\mathbf{y}}}$ is the residential *income* flow shares.

The comparative statics can be expressed as follows.

Theorem 8. *Given a solution* $\mathbf{x} = [\tilde{\mathbf{L}}_{\mathbf{R}}', \tilde{\mathbf{L}}_{\mathbf{F}}', \tilde{\bar{\mathbf{y}}}', \tilde{\mathbf{r}}', \boldsymbol{\upsilon}_F', \tilde{\mathbf{w}}', \boldsymbol{\upsilon}_R']'$ *to the system of equations* (L_R)-(υ_{Ri}), *the elasticities of endogenous variables with respect to commuting costs are given* *by the following expression, whenever the matrix* $\mathbf{A}$ *is full rank:*

$$\frac{\partial \ln \mathbf{x}^*}{\partial \ln \mathbf{d}} = -\mathbf{A}^{-1}\mathbf{T}, \tag{65}$$

where $\mathbf{A} \equiv \overline{\mathbf{A}}\,\tilde{\boldsymbol{\Gamma}}$, $\tilde{\boldsymbol{\Gamma}} \equiv \boldsymbol{\gamma} \otimes \mathbf{I}_N$, *and*

$$\overline{\mathbf{A}} \equiv \mathbf{I}_{7N} - \begin{bmatrix} \boldsymbol{\alpha}_X, \\ \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\upsilon}} \end{bmatrix}.$$

$$\mathbf{T} \equiv \begin{bmatrix} \theta\left(\frac{\mathbf{L}}{\mathbf{L_R}} \otimes \mathbf{1}\right) \circ (\mathbf{1} \otimes \mathbf{I}_N) \\ \\ \theta\left(\mathbf{1} \otimes \frac{\mathbf{L}'}{\mathbf{L_F}}\right) \circ (\mathbf{I}_N \otimes \mathbf{1}) \\ \\ \theta\left(\frac{\mathbf{L} \circ \mathbf{w}'}{\mathbf{L_R} \circ \bar{\mathbf{y}}} \otimes \mathbf{1}\right) \circ (\mathbf{1} \otimes \mathbf{I}_N) \\ \mathbf{0}_{N,N^2} \\ \mathbf{0}_{N,N^2} \\ \mathbf{0}_{N,N^2} \\ \mathbf{0}_{N,N^2} \end{bmatrix}.$$

where $\mathbf{0}$ *is an* N *by* N *matrix of zeros,* $\mathbf{0}_{N,N^2}$ *is an* N *by* N^2 *matrix of zeros,* $\mathbf{1}$ *is a 1 by* N *row vector of ones,* $\mathbf{I}$ *is the identity matrix,* $\circ$ *denotes the Hadamard product, and* $\otimes$ *denotes the Kronecker product.*

$\mathbf{A}$ *is a* $7N \times 7N$ *matrix and* $\mathbf{T}$ *is a* $7N \times N^2$ *matrix.*

We note that only the last equation is written in a non-constant-elasticity form. Therefore, only the elasticities and the flow share matrices are needed for the first 6 row blocks of the matrix $\overline{\mathbf{A}}$. For the last row block, the researcher is also required to use knowledge of the initial equilibrium, through the ratio of commercial and residential floorspace. The dependency on commuting costs d_{ij} only appears in the equations that take the constant-elasticity form, and hence separable between endogenous variables and location-specific parameters. Hence, the $\mathbf{T}$ matrix only requires knowledge of parameters and flow-shares (knoweledge of endogenous variables is not required).