NBER WORKING PAPER SERIES

THE WELFARE EFFECTS OF TRANSPORTATION INFRASTRUCTURE IMPROVEMENTS

Treb Allen
Costas Arkolakis

Working Paper 25487
http://www.nber.org/papers/w25487

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2019, Revised December 2020

We thank our editor Adam Szeidl, discussants Pol Antras and Caitlin Gorback, and four anonymous referees for exceptionally helpful feedback and suggestions. The project also benefited from discussions with Takashi Akamatsu, David Atkin, Don Davis, Dave Donaldson, Gilles Duranton, Pablo Fajgelbaum, Ed Glaeser, Xiangliang Li, Sam Kortum, and Steve Redding. We are grateful to the experts in the Economic Analysis Division at the Volpe National Transportation Systems Center, especially Sari Radin, for their advice and guidance on the project. We thank Albert Chen for exceptional research assistance. The authors acknowledge support by the National Science Foundation under grants SES-1658838 and SES-1658875 and the National Bureau of Economic Research's “Economics of Transportation in the 21st Century” Initiative. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Welfare Effects of Transportation Infrastructure Improvements
Treb Allen and Costas Arkolakis
NBER Working Paper No. 25487
January 2019, Revised December 2020
JEL No. H54,R12,R13,R41,R42

ABSTRACT

Each year in the U.S., hundreds of billions of dollars are spent on transportation infrastructure and billions of hours are lost in traffic. We incorporate traffic congestion into a quantitative general equilibrium spatial framework and apply it to evaluate the welfare impact of transportation infrastructure improvements. Our approach yields analytical expressions for transportation costs between any two locations, the traffic along each link of the transportation network, and the equilibrium distribution of economic activity across the economy, each as a function of the underlying quality of infrastructure and the strength of traffic congestion. We characterize the properties of such an equilibrium and show how the framework can be combined with traffic data to evaluate the impact of improving any segment of the infrastructure network. Applying our framework to both the U.S. highway network and the Seattle road network, we find highly variable returns to investment across different links in the respective transportation networks, highlighting the importance of well-targeted infrastructure investment.

Treb Allen
Department of Economics
Dartmouth College
6106 Rockefeller Hall
Hanover, NH 03755
and NBER
treb@dartmouth.edu

Costas Arkolakis
Department of Economics
Yale University, 28 Hillhouse Avenue
P.O. Box 208268
New Haven, CT 06520-8268
and NBER
costas.arkolakis@yale.edu

1 Introduction

More than a trillion dollars is spent on transportation infrastructure across the world each year (Lefevre, Leipziger, and Raifman, 2014). In the U.S. alone – where annual spending on highways exceeds $150 billion – the average driver spends an average of 42 hours a year in traffic, generating economic losses exceeding these direct costs (ASCE, 2017). Evaluating the impact of infrastructure investments in the presence of such traffic congestion is difficult. On the one hand, improvements to one part of the infrastructure network causes drivers to alter their routes, changing traffic patterns and congestion throughout the network. On the other hand, changes in traffic patterns affects the spatial distribution of economic activity, as individuals re-optimize where to live, work, and/or consume. But as the spatial distribution of economic activity determines the underlying traffic patterns, these two hands are intricately intertwined, resulting in a complex feedback loop between routing, traffic, congestion, and the spatial distribution of economic activity.

We propose a new tractable spatial framework that incorporates traffic congestion and apply it to evaluate the welfare impact of transportation infrastructure improvements. We embed a route choice problem into two spatial models where the cost of traversing a particular link depends on the equilibrium amount of traffic on that link. Our approach yields analytical expressions for transportation costs between any two locations, the traffic along each link of the transportation network, and the equilibrium distribution of economic activity across the economy. We characterize the properties of such an equilibrium, highlighting how the presence of traffic congestion shapes those properties. We then show how the framework can be combined with readily available traffic data to evaluate the welfare impact of improving any segment of the infrastructure network. Finally, we evaluate the welfare impact in two settings: (1) the U.S. highway network; and (2) the Seattle road network. In both cases we find on average positive but highly variable returns to investment, showing the importance of well-targeted infrastructure investment.

Our framework begins with a modest departure from two widely used quantitative general equilibrium models: an economic geography model where agents choose a location to live (as in Allen and Arkolakis (2014)) and engage in trade between locations (as in Eaton and Kortum (2002)), and an urban model where agents choose where to live and where to work within a city (as in Ahlfeldt, Redding, Sturm, and Wolf (2015)). In Eaton and Kortum (2002), it is assumed that while each location has a idiosyncratic productivity for producing each type of good, the transportation technology is identical for all goods. Similarly, in Ahlfeldt, Redding, Sturm, and Wolf (2015), while it is assumed that each individual has idiosyncratic preferences for each home-work pair of locations, all individuals incur the same transportation costs when commuting from home to work. In our framework, we allow for transportation costs in both models to also be subject to idiosyncrasies at the route-level. As a result, simultaneous to their choice of where to purchase goods (in the economic geography model) or where to live and work (in the urban model), agents also choose an optimal route through the transportation network.

This departure allows us to derive an analytical expression for the endogenous transportation costs between all pairs of locations as a function of the transportation network. It also allows us to derive an analytical expression for the equilibrium traffic along a link. This expression takes an appealing "gravity" form, where traffic depends only on the cost of travel along the link and the economic conditions at the beginning and end of the link. Those economic conditions turn out to be the familiar market access terms (see e.g. Anderson and Van Wincoop (2003); Redding and Venables (2004)) – the "inward" market access at the start of the link and the "outward" market access at the end – highlighting the close relationship between equilibrium traffic flows and the equilibrium distribution of economic activity. It is this close relationship that allows us to tractably introduce traffic congestion, which we do so in the spirit Vickrey (1967), by assuming transportation costs of traversing a link depend on both the underlying infrastructure and amount of traffic along the link.

Ultimately, we can express the equilibrium distribution of economic activity solely as a function of the underlying infrastructure matrix, the geographic fundamentals of each location, and four model elasticities, one of which is new (the traffic congestion elasticity) and three of which are not (a trade/commuting elasticity, a productivity externality, and an amenity externality). While the mathematical structure the equilibrium system takes has to our knowledge not been studied before, we prove an equilibrium will exist and provide conditions under which it will be unique. The new mathematical structure also yields new implications: most notably, the presence of traffic congestion implies that the equilibrium is no longer scale invariant. Increasing the size of an economy results in disproportionate changes in bilateral transportation costs due to changes in traffic congestion, reshaping the equilibrium distribution of economic activity.

We then turn to the question of how to apply our framework empirically. We begin by developing a few new tools. First, we derive an analytical relationship between traffic flows along a network and bilateral trade/commuting flows between an origin and destination; in contexts such as our own where we observe both, this serves as a model validation check. Second, we show that the "exact-hat" approach of conducting counterfactuals (see Dekle, Eaton, and Kortum (2008); Costinot and Rodríguez-Clare (2014); Redding and Rossi-Hansberg (2017)) can be applied to our framework, albeit using (readily available) traffic data rather than harder to observe bilateral trade/commuting data. Third, we provide conditions under which one can recover the necessary traffic congestion elasticity from a regression of speed of travel on traffic, where the traffic gravity equation provides guidance in the search for an appropriate instrument for traffic.

Finally, we calculate the welfare impact of transportation infrastructure improvements in two settings: (1) the U.S. highway network (using the economic geography variant of the framework); and (2) the Seattle road network (using the urban variant). In both cases, we begin by showing that the observed network of traffic flows, appropriately inverted through the lens of the model, does a good job predicting the observed matrix of trade and commuting flows, respectively. We then estimate the strength of traffic congestion, finding in both cases substantial traffic congestion. We proceed by estimating the welfare elasticity of improving each link on each road network. We find highly variable elasticities across different links, with the greatest gains in the densest areas of economic activity and at choke-points in the network. Here, traffic congestion plays a particularly important role, as there is only a modest positive correlation between these welfare elasticities and those that one would calculate in a standard model ignoring congestion forces.

Finally, we combine our welfare elasticities with detailed cost estimates of improving each link (which depends on the number of lane-miles needed to be added as well as the geographic topography and the density of economic activity along the link) to construct an estimate of the return on investment for each link. For the U.S. highway network, we estimate an average annual return on investment of 108%; for the Seattle road network that figure is 16%. Both averages, however, belie substantial heterogeneity across links. For the U.S. highway network, the returns on investment for a handful of highways serving as connectors just outside major metropolitan areas exceed 400%; in Seattle, a number of links surrounding downtown have annualized returns exceeding 60%. Conversely, a substantial fraction of U.S. highway links (mainly through the mountain west) and nearly half the links in Seattle are estimated to have a negative return on investment. Taken together, these results highlight the importance of targeting infrastructure improvements to the appropriate locations in the infrastructure network.

The primary contribution of the paper is to develop a quantitative general equilibrium spatial framework that incorporates traffic congestion and can be applied to empirically evaluate the welfare impact of transportation infrastructure improvements. In doing so, we seek to connect two related – but thus far distinct – literatures.

The first literature seeks to understand the impacts of infrastructure improvements on the distribution of economic activity. This literature is mostly the domain of spatial economists; early examples include Fogel (1962, 1964); recent quantitative work on the subject that incorporates rich geographies and general equilibrium linkages across locations include Donaldson

(2018), Allen and Arkolakis (2014), Donaldson and Hornbeck (2016) in an inter-city context and the work of Ahlfeldt, Redding, Sturm, and Wolf (2015) and Tsivanidis (2018) in an intra-city context; Redding and Turner (2015) and Redding and Rossi-Hansberg (2017) offer excellent reviews. While the details of these models vary, a unifying characteristic is that the transportation costs are treated as exogenous model parameters (usually determined by the least cost route, as computed using Dijkstra's algorithm or the "Fast Marching Method" pioneered by Osher and Sethian (1988) and Tsitsiklis (1995)). As a result, this literature abstracts from the effect of infrastructure improvements on how changes in the use of the transportation network affects the transportation costs themselves through traffic congestion.

Relative to this literature, we make two contributions: first, we provide an analytical relationship between the transportation network and the bilateral costs of travel through the network, obviating the need to rely on computational methods. Second (and more importantly), we allow the transportation costs to respond endogenously through traffic congestion to changes in the distribution of economic activity. This force has been identified as empirically relevant (see Duranton and Turner (2011)) but thus far has been absent in such quantitative modeling. Our analysis retains the key analytical benefits of that previous work but also provides a comprehensive framework to analyze the effects of traffic both theoretically and empirically.

The second literature seeks to understand the impacts of infrastructure improvements on the transportation network. This literature is mostly the domain of transportation economics; early examples include Beckmann, McGuire, and Winsten (1955) and seminal textbook of Sheffi (1985); recent work on the subject includes Bell (1995), Akamatsu (1996), De Palma, Kilani, and Lindsey (2005), Eluru, Pinjari, Guo, Sener, Srinivasan, Copperman, and Bhat (2008), Mattsson, Weibull, and Lindberg (2014); Galichon (2016) provides a comprehensive theoretical treatment and Chapter 10 of De Palma, Lindsey, Quinet, and Vickerman (2011) provides an excellent review. While the details of these models vary, a unifying characteristic is that the economic activity at each node in the network is taken as given, so the literature abstracts from how changes in the transportation costs affects this distribution of economic activity.

Relative to this literature, we also make two contributions: first, we provide an analytical solution for the equilibrium traffic along each link in the network that highlights the close relationship between traffic and the equilibrium distribution of economic activity. Second (and more importantly), we allow infrastructure improvements to affect traffic not only through changing route choices (and congestion) on the network, but also through the resulting equilibrium changes in the distribution of economic activity.

Most closely related to this paper is parallel work by Fajgelbaum and Schaal (2020), who characterize the optimal transportation network in a similarly rich geography and also in the presence of traffic congestion. In that important work, the focus is on an efficient equilibrium of a flexible spatial model, as it is assumed that the social planner can implement optimal Pigouvian taxes to offset the externalities created by traffic congestion. Our focus, instead, is on the competitive equilibrium of constant elasticity quantitative spatial models where the presence of traffic congestion (and/or productivity and amenity externalities) means the equilibrium is (generically) inefficient, which in the two settings we consider is more appropriate given the absence of congestion tolls. Relative to Fajgelbaum and Schaal (2020), a separate contribution is that the analytical tractability of the framework developed here facilitates the use of many of the tools developed previously by the quantitative spatial literature, such as the ability to evaluate the welfare impact of infrastructure improvements using readily available traffic data and the use of "exact hat algebra" methodology to compute counterfactuals.[1]

The remainder of the paper proceeds as follows. In the next section, we incorporate the routing choice of agents in economic geography and urban variants of the framework. In Section 3, we provide analytical expressions for the endogenous transportation costs and traffic flows in the presence of traffic congestion. In Section 4, we combine the results of the previous sections to characterize the equilibrium distribution of economic activity and traffic. In Section 5, we develop a set of tools for applying the framework empirically. In Section 6, we implement these tools to examine the welfare impacts of improvements to the U.S. highway network and the Seattle road network. Section 7 concludes.

2 Optimal Routing in Two Spatial Models

In this section, we embed an optimal routing problem into two quantitative spatial models: an economic geography model (where goods are traded between locations subject to trade costs) and an urban model (where workers commute between locations subject to commuting costs). We show that both models yield identical expressions for the endogenous transportation costs, and mathematically identical equilibrium conditions as a function of these costs. This allows us to derive analytical expressions for costs, traffic, and congestion in both frameworks, a task we undertake in Section 3.

[1] The tractability of our approach is evinced by the number of recent working papers who have proposed extensions to it since its original dissemination. These include extending the framework to consider multiple types of transportation networks and transshipment (as in Fan, Lu, and Luo (2019) and Fan and Luo (2020), respectively), extending the framework to include endogenous development of transportation capabilities in locations (as in Ducruet, Juhász, Nagy, Steinwender, et al. (2020)), and extending the framework to multiple sectors with economies of scale in traffic rather than traffic congestion (as in Ganapati, Wong, and Ziv (2020)).

For both models, we posit the following geography. Suppose the economy consists of a finite number of locations $i \in \{1, ..., N\} \equiv \mathcal{N}$ arrayed on a network and inhabited by $\bar{L}$ individuals. Mathematically, this networks is represented by an $N \times N$ matrix $\mathbf{T} = [t_{kl} \geq 1]$, where t_{kl} indicates the (ad valorem) cost incurred from moving *directly* from k to l along a *link* (if no link between k and l exists, then $t_{kl} = \infty$).[2] We refer to $\mathbf{T}$ as the *transportation network* and emphasize that it is endogenous and will depend on the equilibrium traffic congestion.

Moving goods (in the economic geography model) or people (in the commuting model) from an origin i to a destination j entails taking a *route* r through the network. Mathematically, r is a sequence of locations beginning with location i and ending with location j, i.e. $r \equiv \{i = r_0, r_1, ..., r_K = j\}$, where K is the number of links crossed on the route, i.e. the *length* of route r. Because iceberg costs are multiplicative, the total costs incurred from moving from i to j along route r of length K is then $\prod_{k=1}^{K} t_{r_{k-1},r_k}$.[3] Let $\Re_{ij}$ denote the set of all the (countably infinite) possible routes from i to j.

2.1 An Economic Geography Model with Optimal Routing

We first embed a routing framework into an economic geography model where goods are traded across locations and labor is mobile, as in Allen and Arkolakis (2014) and Redding (2016). In what follows, unless otherwise noted, we refer the interested reader to Online Appendix C for detailed derivations.

2.1.1 Setup

An individual residing in location i supplies her endowed unit of labor inelastically for the production and shipment of goods, for which she receives a wage w_i and from which she purchases quantities of a continuum of consumption goods $\nu \in [0, 1]$ with constant elasticity of substitution (CES) preferences with elasticity of substitution $\sigma \geq 0$. Labor is the only factor used in the production and shipment of goods and income and the corresponding wage is denoted by w_j. Let Y^W and $\bar{L}$ denote the world income and world labor endowment, respectively; in what follows, we choose world per-capita income as our numeraire, i.e. $Y^W/\bar{L} = 1$,

[2] Following the literature on graph theory (see e.g. p.14 of Szabo (2015) or p.218 from Chartrand (1977)), we assume that $t_{ii} = \infty$ to exclude self loops; however, below we allow agents in i to choose the "null" path (which is the only admissible path of length 0) where they source goods / work where they reside, thereby incurring no transportation costs.

[3] We follow the tradition of the spatial literature by treating transportation costs as ad valorem (iceberg). In Online Appendix D.1, we consider an alternative framework where costs incurred traveling through the network are additive and show that one can derive a similar expression for the endogenous transportation costs below.

which implies that the value of trade is measured in average units of labor.

Each location $i \in \mathcal{N}$ is endowed with a constant returns to scale technology for producing and shipping each good $\nu \in [0,1]$ to each destination $j \in \mathcal{N}$ along each route $r \in \Re_{ij}$, which under perfect competition yields the following price of good $\nu \in [0,1]$ in destination $j \in \mathcal{N}$ from origin $i \in \mathcal{N}$ along route $r \in \Re_{ij}$:

$$p_{ij,r}(\nu) = w_i \frac{\prod_{k=1}^{K} t_{r_{k-1},r_k}}{\varepsilon_{ij,r}(\nu)}.$$

Following Eaton and Kortum (2002), we assume $\varepsilon_{ij,r}(\nu)$ is independently and identically Frechet distributed across routes and goods distributed with scale parameter $1/A_i$ and shape parameter θ. Individuals purchase each good $\nu \in [0,1]$ from the cheapest source (i.e. location-route).

While the basic setting is similar to Eaton and Kortum (2002), the innovation here is that individuals choose both a location and route to source each good (rather than just a location). But why would a consumer not simply choose to purchase the goods from the cheapest source along the least cost route? Some of the value of this choice of modeling arises from the great tractability it yields below. Yet this added "noise" is also is plausible in the presence of traffic congestion, as there will be many alternative routes that yield approximately the same costs.[4] If all consumers were to use the least cost route, then infinitesimal deviations from Mogridge's hypothesis would result in large changes in agents' route choice; empirically, an infinitely elastic route choice is unrealistic; theoretically, it would lead to a nightmare of corner solutions (as e.g. noted by Eaton and Kortum (2012) as original impetus for the Eaton and Kortum (2002) framework).[5]

A related concern is with the assumption that agents simultaneously choose the location that sources the good and the route over which it is supplied. Should agents not first choose where to purchase a good and then decide how to ship it? It turns out the timing assumption is not crucial: one can construct a model with just such a timing assumption that is formally isomorphic to the framework presented here (see Online Appendix D.2). Instead, what is enormously helpful (and which the simultaneous choice over locations and routes ensures) is that agents' demand elasticities for location and route are the same. Deviations from this assumption – while computationally straightforward – come so at the loss of substantial

[4] This is known as Modgridge's hypothesis, quoted as originally stating "For trip origins at any particular distance from the center of London, peak hour journey times by car and rail to central destinations are equal" (Holden, 1989).

[5] In practice, the addition of noise is of little consequence when calculating transportation costs. In the empirical contexts considered below, the correlation between the (log) transportation costs estimated with noise and the (log) transportation costs along the least cost route exceeds 0.99 (for the U.S. highways) and 0.98 (for the Seattle road network).

analytical tractability and ensuing economic insight.[6]

2.1.2 An analytical expression for transportation costs

We now characterize the fraction and value of goods shipped on each route between each origin and destination. The probability that $j \in \mathcal{N}$ purchases good $\nu \in [0,1]$ from $i \in \mathcal{N}$ along route $r \in \Re_{ij}$, $\pi_{ij,r}$, can be written as:

$$\pi_{ij,r} = \frac{(w_i/A_i)^{-\theta}\left(\prod_{l=1}^{K} t_{r_{l-1},r_l}^{-\theta}\right)}{\sum_{k\in\mathcal{N}}(w_k/A_k)^{-\theta}\sum_{r'\in\Re_{kj}}\prod_{l=1}^{K} t_{r'_{l-1},r'_l}^{-\theta}}. \tag{1}$$

To determine the total value of goods shipped from $i \in \mathcal{N}$ to $j \in \mathcal{N}$, X_{ij}, we sum across all routes, recalling from Eaton and Kortum (2002) that the expenditure shares are equal to the probability of purchasing a good:

$$X_{ij} = \sum_{r\in\Re_{ij}} \pi_{ij,r} E_j = \frac{\tau_{ij}^{-\theta}(w_i/A_i)^{-\theta}}{\sum_{k\in\mathcal{N}}\tau_{kj}^{-\theta}(w_k/A_k)^{-\theta}} E_j, \tag{2}$$

where:

$$\tau_{ij} \equiv \left(\sum_{r\in\Re_{ij}}\left(\prod_{l=1}^{K} t_{r_{l-1},r_l}^{-\theta}\right)\right)^{-\frac{1}{\theta}} \tag{3}$$

is the *transportation costs* from i to j. Note that expression (2) is identical to that of Eaton and Kortum (2002); however, rather than the transportation cost τ_{ij} being taken as given, here it is determined by the least cost routing problem through the (endogenous) transportation network.

2.1.3 Market Access and Gravity

While (2) provides an analytical expression for the value of bilateral trade flows, it turns out it is convenient for what follows to express it in market access terms. To do so, we first impose two equilibrium market clearing conditions: (1) total income Y_i in each location is is equal to its total sales; and (2) total expenditure E_i in each location is equal to its total purchases:

$$Y_i = \sum_{j=1}^{N} X_{ij},\ \ E_i = \sum_{j=1}^{N} X_{ji}. \tag{4}$$

[6] There are two places where an equal demand elasticity for location and route greatly increases the tractability: first, in transforming the equilibrium conditions of the model written as a function of transportation costs to a function of the transportation matrix (where it allows for a linear inversion); second, in deriving the traffic gravity equation (where it allows for an explicit rather than implicit analytical form.)

We can then follow Anderson and Van Wincoop (2003) and Redding and Venables (2004) to re-write the gravity equation (2) as follows:

$$X_{ij} = \tau_{ij}^{-\theta} \times \frac{Y_i}{\Pi_i^{-\theta}} \times \frac{E_j}{P_j^{-\theta}}, \tag{5}$$

where Π_i is a producer price index capturing the (inverse) of producer market access:

$$\Pi_i \equiv \left(\sum_{j=1}^{N} \tau_{ij}^{-\theta} E_j P_j^{\theta} \right)^{-\frac{1}{\theta}} = A_i L_i Y_i^{-\frac{\theta+1}{\theta}}, \tag{6}$$

and P_j is the consumer price index capturing the (inverse) of consumer market access:

$$P_j = \left(\sum_{i=1}^{N} \tau_{ij}^{-\theta} Y_i \Pi_i^{\theta} \right)^{-\frac{1}{\theta}}. \tag{7}$$

A lower value of P_j indicates that consumers in location i have greater access to producers in other markets, and a lower value of Π_i indicates that producers have greater access to consumers in other markets.

2.1.4 Equilibrium

Finally, we calculate the equilibrium distribution of population and economic output across space. Following Allen and Arkolakis (2014), we write the welfare of residents in location $j \in \mathcal{N}$, W_j, as:

$$W_j = \frac{w_j}{P_j} u_j, \tag{8}$$

where u_j is an amenity value of living in location $j \in \mathcal{N}$. We assume that there is free labor mobility across locations and we focus in equilibria where welfare equalizes across locations, $W_j = \bar{W}$, and every location is populated.[7]

We further allow for the possibility that productivities and amenities potentially depend on the measure of workers in a given location as follows:

$$A_i = \bar{A}_i L_i^{\alpha},\ u_i = \bar{u}_i L_i^{\beta}, \tag{9}$$

where $\bar{A}_i > 0$ and $\bar{u}_i > 0$ are the *local geography* of productivity and amenities, and $\alpha, \beta \in \mathbb{R}$

[7] This assumption, combined with congestion spillovers introduced later, simply introduces a labor supply function that increases in the real wage offered in a location. Various microfoundations of such a labor supply function have been discussed in the literature, see for example Allen and Arkolakis (2014); Redding (2016); Redding and Rossi-Hansberg (2017); Allen, Arkolakis, and Takahashi (2020).

govern the strength of the productivity and amenity externalities, respectively.[8]

Combining the definitions in (9), equation (5), the market clearing conditions (4), imposing balanced trade (i.e. $E_i = Y_i$) and welfare equalization (i.e. condition (8)), we obtain the following equilibrium conditions:

$$\bar{A}_i^{-\theta} y_i^{1+\theta} l_i^{-\theta(1+\alpha)} = \chi \sum_{j=1}^{N} \tau_{ij}^{-\theta} \bar{u}_j^{\theta} y_j^{1+\theta} l_j^{\theta(\beta-1)} \tag{10}$$

$$\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)} = \chi \sum_{j=1}^{N} \tau_{ji}^{-\theta} \bar{A}_j^{\theta} y_j^{-\theta} l_j^{\theta(\alpha+1)} \tag{11}$$

where $y_i \equiv Y_i/Y^W$ and $l_i \equiv L_i/\bar{L}$ are the share of world income and world labor in location $i \in \mathcal{N}$, respectively, and $\chi \equiv \left(\frac{\bar{L}^{(\alpha+\beta)}}{\bar{W}}\right)^{\theta}$ is an endogenous scalar capturing the (inverse) of the equilibrium welfare of the system.[9] Given productivities $\{A_i\}$, amenities $\{u_i\}$, and transportation costs $\{\tau_{ij}\}$, the $2N$ equations (10) and (11) can be solved for the $2N$ equilibrium shares of income $\{y_i\}$ and labor $\{l_i\}$ in all locations. It is essential to note that the transportation costs themselves are endogenous and – through traffic congestion – will respond to the equilibrium distribution of economic activity; hence, these conditions only provide part of the story. We address the remainder of the story below in Section 3. First, however, we turn to another spatial model.

2.2 An Urban Model with Optimal Routing

We next embed a routing framework in an urban model where agents commute between their place of residence and their place of work, as in Ahlfeldt, Redding, Sturm, and Wolf (2015).

2.2.1 Setup

An individual $\nu \in [0,1]$ residing in city block $i \in \mathcal{N}$ who works in city block $j \in \mathcal{N}$ and commutes via route r of length K to work receives a payoff $V_{ij,r}(\nu)$ that depends on (1)

[8] As noted in Allen and Arkolakis (2014), the presence of productivity and amenity spillovers create formal isomorphisms between a large set of economic geography models and also play an important role in determining the qualitative and quantitative implications of the model. We will contrast the implications of these (now standard) spillovers to the (new) traffic congestion spillovers below.

[9] That χ combines both the equilibrium welfare $\bar{W}$ and the aggregate population $\bar{L}$ demonstrates that the whether one treats the economy as "closed" (so $\bar{L}$ is fixed and $\bar{W}$ is endogenous) or "open" (so that $\bar{W}$ is fixed and $\bar{L}$ is endogenous) has no bearing on the equilibrium distribution of economic activity $\{l_i, y_i\}_{i\in\mathcal{N}}$ nor on the value that χ takes, i.e. χ is a sufficient statistic for welfare in either scenario. This is closely related to the fact that, conditional on transportation costs, the equilibrium is scale invariant – i.e. changes in $\bar{L}$ have no effects on χ or the equilibrium distribution of economic activity – a point we discuss in detail in Section 4.3.

the wage in the workplace, w_j; (2) the amenity value of residence, u_i; (3) the time spent commuting; and (4) an idiosyncratic (Frechet distributed with shape parameter θ) route-, origin-, and destination-specific term:

$$V_{ij,r}(\nu) = \left(u_i w_j / \prod_{l=1}^{K} t_{r_{l-1},r_l} \right) \times \varepsilon_{ij,r}(v) .$$

Individual ν chooses where to live, work, and which route to take in order to maximize $V_{ij,r}(\nu)$. That is, we extend the framework of Ahlfeldt, Redding, Sturm, and Wolf (2015) to introduce heterogeneity across individuals in their preference not only of where to live and work but also of what route to take when commuting between the two. Like in the economic geography framework above, this additional "noise" both substantially increase the tractability and generates an empirically plausible finite elasticity to the costs of different routes between home and work. And as above, the assumption that the three choices of where to live, where to work, and what route to take share the same elasticity – while straightforward to relax – greatly facilitate the tractability of the derivations and ensuing economic insight that follows.

We assume each location j produces a homogeneous and costlessly traded good with a constant returns to scale production function where labor is the only factor of production with productivity A_j. Taking the price of the good as the numeraire, this implies that the equilibrium real wage is the marginal product of labor $w_j = A_j$.

2.2.2 An analytical expression for transportation costs

The probability a worker chooses to live in i, work in j, and commute via route r can be written as:

$$\pi_{ij,r} = \frac{\prod_{l=1}^{K} t_{r_{l-1},r_l}^{-\theta} \times u_i^{\theta} \times w_j^{\theta}}{\sum_{i,j} \prod_{l=1}^{K} t_{r_{l-1},r_l}^{-\theta} \times u_i^{\theta} \times w_j^{\theta}}, \tag{12}$$

where we re-use the notation from the economic geography model for reasons that will become apparent below. This implies that the total number of workers residing in i and working in j, L_{ij},can then be determined by simply summing across all routes and multiplying by the aggregate population $\bar{L}$, yielding for all $i \in \mathcal{N}$and $j \in \mathcal{N}$:

$$L_{ij} = \sum_{r \in \Re_{ij}} L_{ij,r} = \tau_{ij}^{-\theta} \times u_i^{\theta} \times w_j^{\theta} \times \frac{\bar{L}}{\bar{W}^{\theta}}, \tag{13}$$

where transportation costs τ_{ij} are given again by (3) and $\bar{W} \equiv E\left[\max_{i,j,r} V_{ij,r}(\nu)\right] = \left(\sum_{ij} \tau_{ij}^{-\theta} \times u_i^{\theta} \times w_j^{\theta}\right)^{\frac{1}{\theta}}$ is the expected welfare of a resident in the city.

2.2.3 Market Access and Gravity

As in the trade model, we can express the gravity commuting equation (13) in market access terms. To do so, we impose the following two market clearing conditions: (1) we require that the total number of residents in i, L_i^R, is equal to the commuting flow to all workplaces; and (2) we require that the total number of workers in j, L_j^F, is equal to the commuting flow from all residences:

$$L_i^R \equiv \sum_j L_{ij},\ L_j^F \equiv \sum_i L_{ij}. \tag{14}$$

We can write the gravity commuting equation (13) as follows:

$$L_{ij} = \tau_{ij}^{-\theta} \times \frac{L_i^R}{\Pi_i^{-\theta}} \times \frac{L_j^F}{P_j^{-\theta}}, \tag{15}$$

where Π_i is a resident price index capturing the (inverse of) the commuting market access residents in i have to firms in all locations:

$$\Pi_i = \left(\sum_j \tau_{ij}^{-\theta} L_j^F P_j^{\theta}\right)^{-\frac{1}{\theta}} = u_i \left(L_i^R\right)^{-\frac{1}{\theta}} \left(\frac{\bar{L}}{\bar{W}^{\theta}}\right)^{\frac{1}{2\theta}}, \tag{16}$$

and P_i is a firm price index capturing the (inverse of) the commuting market access firms in j have to residents in all locations:

$$P_j = \left(\sum_j \tau_{ij}^{-\theta} L_i^R \Pi_i^{\theta}\right)^{-\frac{1}{\theta}} = w_j \left(L_j^F\right)^{-\frac{1}{\theta}} \left(\frac{\bar{L}}{\bar{W}^{\theta}}\right)^{\frac{1}{2\theta}}. \tag{17}$$

Note that we re-use the notation from the economic geography framework above: both models $\Pi_i^{-\theta}$ captures the "outward" market access and $P_j^{-\theta}$ captures the "inward" market access with respect to the flows from i to j.

2.2.4 Equilibrium

As in the economic geography model, we assume that productivities and amenities are affected by commercial and residential population, respectively, as follows:

$$A_i = \bar{A}_i \left(L_i^F\right)^{\alpha},\; u_i = \bar{u}_i \left(L_i^R\right)^{\beta}, \tag{18}$$

where $\bar{A}_i > 0$ and $\bar{u}_i > 0$ are again the *fundamental* components of productivity and amenities and α, β the respective elasticities.

Substituting equations (18) into the commuting gravity equation (13) and imposing the equilibrium market clearing conditions (14) yields the following system of equations:

$$\left(l_i^R\right)^{-\theta\beta+1} = \chi \sum_j \tau_{ij}^{-\theta} \bar{u}_i^{\theta} \bar{A}_j^{\theta} \left(l_j^F\right)^{\alpha\theta} \tag{19}$$

$$\left(l_i^F\right)^{-\theta\alpha+1} = \chi \sum_j \tau_{ji}^{-\theta} \bar{u}_j^{\theta} \bar{A}_i^{\theta} \left(l_j^R\right)^{\beta\theta}, \tag{20}$$

where $l_i^R \equiv L_i^R/\bar{L}$ and $l_i^F \equiv L_i^F/\bar{L}$ are the share of workers living and working, respectively, in location i and $\chi \equiv \left(\frac{\bar{L}^{(\alpha+\beta)}}{\bar{W}}\right)^{\theta}$ is again the (inverse) of the equilibrium welfare of the system. As in the trade model above, given transportation costs $\{\tau_{ij}\}$, productivities $\{A_i\}$, and amenities $\{u_i\}$, equations (19) and (20) can be solved to determine the equilibrium distribution of where people live $\left\{l_i^R\right\}$ and where they work $\left\{l_i^F\right\}$. Once again, however, the transportation costs themselves are endogenously determined and will respond to the distribution of economic activity through traffic congestion.

2.3 Taking Stock: Gravity and Optimal Routing on the Network

We now compare the economic geography and urban models. As is evident, the two setups are very similar, sharing (1) identical expressions for the (endogenous) bilateral trade/commuting costs (summarized in equation (3)); (2) identical gravity expressions for the bilateral flow of goods / commuters as a function of bilateral costs and market access (summarized in equations (5) and (13), respectively); and (3) and mathematically equivalent equilibrium conditions (summarized in equations (10) and (11) for the economic geography model and equations (19) and (20) for the urban model). Indeed, the only distinction between the two models is the particular log linear relationship between market access variables $\Pi_i^{-\theta}$ and $P_j^{-\theta}$ and the equilibrium economic activity in the origin (Y_i and L_i^R, respectively) and the destination (E_j and L_j^F, respectively): the equilibrium conditions in both models as functions of the market access variables and economic activities are identical.[10] These similarities allow us to introduce endogenous transportation costs through equilibrium traffic congestion in both frameworks using a unified set of tools, which we turn to next.

3 Transportation Costs, Traffic, and Congestion

In this section, we provide analytical solutions for the equilibrium transportation costs, traffic, and congestion throughout the infrastructure network.

3.1 Transportation Costs

Both the economic geography and urban models yield transportation costs of the form given in equation (3). By explicitly enumerating all possible routes, equation (3) can be written in matrix notation as follows:[11]

$$\tau_{ij}^{-\theta} = \sum_{K=0}^{\infty} A_{ij}^{K},$$

where $\mathbf{A} \equiv \left[t_{ij}^{-\theta}\right]$, i.e. $\mathbf{A}$ is an $N \times N$ matrix with (i,j) element $t_{ij}^{-\theta}$ (not to be confused with the vector of productivities) and $\mathbf{A}^K = \left[A_{ij}^{K}\right]$, i.e. A_{ij}^{K} is the (i,j) element of the matrix $\mathbf{A}$ to the matrix power K.[12] As in Bell (1995), as long as the spectral radius of $\mathbf{A}$ is less than one, the geometric sum can be expressed as:[13]

$$\sum_{K=0}^{\infty} \mathbf{A}^K = (\mathbf{I} - \mathbf{A})^{-1} \equiv \mathbf{B},$$

where $\mathbf{B} = [b_{ij}]$ is simply the Leontief inverse of the weighted adjacency matrix. As a result, the transportation cost from i to j can be written as a simple function of the infrastructure

[10] That our model yields a log-linear relationship between local economic outcomes and market access terms means that it generates a structural interpretation to the empirical specification used by a recent literature to estimate the effects of transportation where economic outcomes are projected onto market access terms (see Donaldson (2015) and Redding and Turner (2015) for excellent reviews).

[11] See Appendix A.1 for a detailed derivation.

[12] By summing over all possible routes, there is a direct analogy to the integral formulation of quantum mechanics, which considers all possible paths of the system in between the initial and final states, including those that are absurd by classical standards. Note that while it is straightforward to truncate the summation up to some finite K to restrict consideration to only routes that are not "too" long, doing so would entail a substantial loss of analytical tractability. In the empirical exercises below, the inclusion of more indirect routes is not quantitatively important, as they are chosen with extremely small probability.

[13] A sufficient condition for the spectral radius being less than one is if $\sum_j t_{ij}^{-\theta} < 1$ for all i. The condition will hold if either transportation costs, t_{ij}, between connected locations are sufficiently large, the adjacency matrix is sufficiently sparse (i.e. many locations are not directly connected so that $t_{ij} = +\infty$), or the heterogeneity in preferences across routes is sufficiently small (i.e. θ is sufficiently large).

matrix:

$$\tau_{ij} = b_{ij}^{-\frac{1}{\theta}}. \tag{21}$$

Equation (21) provides an analytical relationship between the transportation network $\mathbf{T} \equiv [t_{kl}]$ and the resulting transportation costs $\{\tau_{ij}\}_{i,j\in\mathcal{N}^2}$, accounting for the choice of the least cost route.

Notice that in the limit case of no heterogeneity ($\theta \to \infty$), the transportation costs converge to those of the least cost route, which is typically solved computationally using the Dijkstra algorithm (see e.g. Donaldson (2018)). Our formulation results in an analytical solution by extending the idiosyncratic heterogeneity already assumed in spatial models to also incorporate heterogeneity over the route chosen. In doing so, our setup bears resemblance to stochastic path-assignment methods used in transportation and computer science literature (c.f. Bell (1995); Akamatsu (1996)); here, however, the endogenous transportation costs arise from –and are determined simultaneously with– a larger general equilibrium spatial model.[14]

3.2 Traffic Flows

We next characterize traffic along a particular link in the infrastructure matrix. This will allow us to introduce traffic congestion into the framework and relate it to observed measures of economic activity. We refer the reader to Appendix A.2 for detailed derivations of the results that follow.

To begin, we characterize the expected number of times in which link (k,l) is used in trade between (i,j), π_{ij}^{kl}, which we refer to as the *link intensity*. We sum across all routes from i to j the product of the probability a particular route is used (conditional on purchasing a product from i to j) and the number of times that route passes through link (k,l), n_r^{kl} (as some routes may use a link more than once):

$$\pi_{ij}^{kl} \equiv \sum_{r\in\Re_{ij}} \left(\frac{\pi_{ij,r}}{\sum_{r'\in\Re_{ij}} \pi_{ij,r'}} \right) n_r^{kl}. \tag{22}$$

Note that for any route r of length K that travels through link (k,l) at least once, there must exist some length $B \in [1, 2, ..., K-1]$ at which the route arrives at link (k,l). As a result, we can calculate π_{ij}^{kl} by explicitly enumerating all possible routes from i to k of length

[14]While equation (21) offers an explicit analytical relationship between the transportation network and the resulting transportation costs that is unavailable with Dijkstra algorithm, in terms of computation, the two share the same operational complexity of $O\left(N^2 \log N\right)$. In practice, however, we find equation (21) offers significant computational advantage over the Djikstra algorithm. For example, in the interstate highway network constructed below ($N = 228$), calculating all bilateral transportation costs takes 0.04 seconds using equation (21) and 116.5 seconds using Djikstra's algorithm – a three orders of magnitude improvement.

B and all possible routes from l to j of length $K-B-1$, which can be expressed as elements of matrix powers of $\mathbf{A}$:

$$\pi_{ij}^{kl} = \tau_{ij}^{\theta} \sum_{K=0}^{\infty} \sum_{B=0}^{K-1} A_{ik}^{B} \times a_{kl} \times A_{lj}^{K-B-1},$$

which, with some matrix calculus, becomes:

$$\pi_{ij}^{kl} = \left(\frac{\tau_{ij}}{\tau_{ik} t_{kl} \tau_{lj}} \right)^{\theta}. \tag{23}$$

This expression – which resembles the one of Akamatsu (1996) derived using an exponential distribution – has a simple intuition: the more "out of the way" the transportation link (k, l) is from the optimal path between i and j (and hence the greater the cost of traveling through link (k, l) along the way from i to j relative to the unconstrained cost of traveling from i to j) the less frequently that link is used.

We now use the above derivation to characterize equilibrium traffic flows along each link of the network. Let Ξ_{kl} be the total traffic over link (k, l), by which we mean the total world value of goods shipped (in the economic geography model) or the total number of commuters (in the urban model) over the link (k, l).[15] To calculate Ξ_{kl}, we sum across all origins, destinations, and routes which travel over link kl, which can be written as:

$$\Xi_{kl} \equiv \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \sum_{r \in \Re_{ij}} \pi_{ij,r} n_r^{kl} E_j = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \pi_{ij}^{kl} X_{ij}$$
$$\Xi_{kl} \equiv \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \sum_{r \in \Re_{ij}} \pi_{ij,r} n_r^{kl} \bar{L} = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \pi_{ij}^{kl} L_{ij},$$

in the economic geography and urban models, respectively. In either case, combining the market access gravity equation ((5) in the economic geography model or (15) in the urban model) with the link intensity equation (23), we obtain the following expression for equilibrium traffic flows:

$$\Xi_{kl} = t_{kl}^{-\theta} \times P_k^{-\theta} \times \Pi_l^{-\theta}. \tag{24}$$

There are several implications of equations (24). First, traffic flows a long a link (k, l) follow a gravity equation, where all determinants of the flow of traffic along link (k, l) are fully

[15] Recall that our choice of world per-capita income as the numeraire means that the total world value of goods shipped over a link is also the average number of persons flowing over a link; that is, traffic in both the economic geography and urban models is measured in persons traveling over a link.

summarized by the cost to travel along the link (t_{kl}) and the economic conditions at the beginning and end of the link. Second, there is a tight connection between the gravity equation for traffic and trade/commuting flows, as the variables summarizing the economic conditions for the traffic gravity equation are the same market access terms P_k and Π_k that shape the economic conditions in the origin and destination in the economic geography and urban models. Third, the intuition for the role that the market access terms play in the traffic gravity equation is straightforward: the greater the inward market access $\left(P_k^{-\theta}\right)$, the more traffic that flows into a link k, and the greater the outward market access $\left(\Pi_l^{-\theta}\right)$, the more traffic that flows out of link l.[16]

Equation (24) takes the cost of traveling along a link t_{kl} as given – we now introduce traffic congestion by a parametric relationship between this cost and the traffic along the link.

3.3 Traffic Congestion

To complete our modeling of traffic flows, we now suppose that the direct cost of traveling over a particular link depends in part on the total traffic flowing over that link through traffic congestion. In particular, we assume that the direct cost of traveling over a link, t_{kl}, depends in part on the amount of traffic over that link Ξ_{kl} through the following simple functional form:

$$t_{kl} = \bar{t}_{kl} \left(\Xi_{kl}\right)^{\lambda}, \tag{25}$$

where $\lambda > 0$ governs the strength of traffic congestion and $\bar{\mathbf{T}} \equiv [\bar{t}_{kl}]$ is the *infrastructure network*. Intuitively, if $\lambda > 0$, the greater the fraction of total economic activity that passes through a link, the more costly traversing that link is. Like the amenity and productivity externalities in equations (9) and (18), the choice of the functional form of equation (25) succinctly allows for transportation costs to depend on an exogenous component (the infrastructure network) and an endogenous component (traffic), with a single structural parameter (λ) governing the relative strength of the two. And like with the amenity and productivity externalities, it has the unattractive feature that the transportation costs is equal to zero when the endogenous component (traffic) is equal to zero. Just as with the amenity and productivity externalities, however, this never occurs in equilibrium, as all agents' idiosyncratic preferences over routes ensures there will be strictly positive traffic on all links. An

[16]In both the economic geography and urban models, traffic flows from an origin to a destination. This abstracts from back-hauling (in the economic geography model) and return commutes (in the urban model). For this reason (and because our traffic data does not indicate a direction of travel), in the empirical exercises below, we consider symmetric improvements to both directions of travel on a given link in the infrastructure network.

additional attractive feature of equation (25) is that can be derived from a simple micro-foundation (presented in Section 5.3) where transportation costs are log-linear functions of travel time and speed is a log-linear function of traffic congestion.

Combining equation (25) with the gravity equation for Ξ_{kl} from equation (24) we immediately obtain:

$$t_{kl} = \bar{t}_{kl}^{\frac{1}{1+\theta\lambda}} \times P_k^{-\frac{\theta\lambda}{1+\theta\lambda}} \times \Pi_l^{-\frac{\theta\lambda}{1+\theta\lambda}} \tag{26}$$

$$\Xi_{kl} = \bar{t}_{kl}^{-\frac{\theta}{1+\theta\lambda}} \times P_k^{-\frac{\theta}{1+\theta\lambda}} \times \Pi_l^{-\frac{\theta}{1+\theta\lambda}} \tag{27}$$

Equation (26) shows how the distribution of economic activity affects transportation costs through traffic congestion. It says that the cost of transiting a link t_{kl} is higher the better the inward market access (lower P_k) at the beginning of the link and/or the better the outward market access (lower Π_l at the end of the link), as both increase traffic along the link, with λ governing the strength of the forces. Equation (27) – which provides the basis for estimating the strength of traffic congestion below – shows traffic flows retain a gravity structure in the presence of traffic congestion. It also highlights that improvements in infrastructure quality endogenously increases the traffic demand for the infrastructure with an elasticity $\frac{\partial \ln \Xi_{kl}}{\partial \ln \bar{t}_{kl}} = -\frac{\theta}{1+\theta\lambda}$, a fact highlighted by Duranton and Turner (2011), and a point we return to in Section 5.3.

4 Traffic Congestion in the Spatial Economy

In Section 2, we characterized the equilibrium distribution of economic activity given transportation costs. In Section 3, we characterized the equilibrium transportation costs given the distribution of economic activity. In this section, we characterize both simultaneously as a function of the fundamental infrastructure network. As noted above, we refer the interested reader to Online Appendix C for detailed derivations.

4.1 Equilibrium

We begin by formally defining our equilibrium: Given a local geography $\left\{\bar{A}_i, \bar{u}_i\right\}_{i\in\mathcal{N}}$, an aggregate labor endowment $\bar{L}$, an infrastructure network $\bar{\mathbf{T}} \equiv [\bar{t}_{kl}]$, and model parameters $\{\alpha, \beta, \theta, \lambda\}$, we define an *equilibrium* to be a distribution of economic activity $\{y_i, l_i\}_{i\in\mathcal{N}}$ in the economic geography model and $\left\{l_i^F, l_i^R\right\}_{i\in\mathcal{N}}$ in the urban model and an aggregate (inverse) welfare $\chi > 0$ such that:

1. Given equilibrium transportation costs $\{\tau_{ij}\}_{i,j\in\mathcal{N}^2}$, the equilibrium distribution of eco-

nomic activity ensures markets clear, i.e. equations (10) and (11) hold in the economic geography model and equations (19) and (20) hold in the urban model;

2. Given the equilibrium transportation network $\mathbf{T} \equiv [t_{kl}]$, agents optimally choose their routes through the network, i.e. equilibrium transportation costs are determined by equation (21); and

3. Given the equilibrium distribution of economic activity, the infrastructure network $\bar{\mathbf{T}} \equiv [\bar{t}_{kl}]$, and agents' optimal route choice, the equilibrium transportation network $\mathbf{T} \equiv [t_{kl}]$ is determined by the equilibrium levels of traffic congestion, i.e. equation (26) holds.

We further define a *strictly positive equilibrium* to be one where the distribution of economic activity is strictly greater than zero in all locations, i.e. $y_i > 0$ and $l_i > 0$ for all $i \in \mathcal{N}$ in an economic geography model and $l_i^F > 0$ and $l_i^R > 0$ for all $i \in \mathcal{N}$ in an urban model. While the first equilibrium condition – market clearing given transportation costs – is standard to all general equilibrium spatial models, the second and third conditions are new, introducing optimal routing on the part of agents and endogenous traffic congestion, respectively. Despite the added complexity of the system, however, it turns out that the equilibrium of the system remains surprisingly tractable. To demonstrate this, we present the derivations for the economic geography model; given their identical mathematical structure, the traffic model calculations proceed similarly.

Recall that equations (10) and (11) characterize the equilibrium distribution of population and income as a function of the endogenous transportation costs $\{\tau_{ij}\}$, i.e. they satisfy equilibrium condition 1. To satisfy equilibrium condition 2, we substitute in equation (21) for the endogenous transportation costs and perform a matrix inversion to re-write the equilibrium conditions as a functions of the infrastructure network rather than the transportation costs:

$$\bar{A}_i^{-\theta} y_i^{1+\theta} l_i^{-\theta(1+\alpha)} = \chi \bar{u}_i^{\theta} y_i^{1+\theta} l_i^{\theta(\beta-1)} + \sum_{j=1}^{N} t_{ij}^{-\theta} \bar{A}_j^{-\theta} y_j^{1+\theta} l_j^{-\theta(1+\alpha)}$$

$$\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)} = \chi \bar{A}_i^{\theta} y_i^{-\theta} l_i^{\theta(\alpha+1)} + \sum_{j=1}^{N} t_{ji}^{-\theta} \bar{u}_j^{-\theta} y_j^{-\theta} l_j^{\theta(1-\beta)}.$$

To incorporate endogenous traffic congestion – satisfying equilibrium condition 3 – we then substitute the endogenous transport costs using equations (24),(26), (27), yielding:

$$y_i^{\frac{1+\theta+\theta\lambda}{1+\theta\lambda}} l_i^{-\frac{\theta(1+\alpha+(\alpha+\beta)\theta\lambda)}{1+\theta\lambda}} = \chi \bar{u}_i^{\theta} \bar{A}_i^{\theta} y_i^{\frac{1+\theta+\theta\lambda}{1+\theta\lambda}} l_i^{\frac{\theta(\beta-1)}{1+\theta\lambda}} + \chi^{\frac{\theta\lambda}{1+\theta\lambda}} \sum_{j=1}^{N} \left(\bar{L}^{\lambda}\bar{t}_{ij}\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{A}_i^{\theta} \bar{u}_i^{\theta \frac{\theta\lambda}{1+\theta\lambda}} \bar{A}_j^{-\frac{\theta}{1+\theta\lambda}} y_j^{\frac{1+\theta}{1+\theta\lambda}} l_j^{-\frac{\theta(1+\alpha)}{1+\theta\lambda}} \tag{28}$$

$$y_i^{\frac{-\theta(1-\lambda)}{1+\theta\lambda}} l_i^{\frac{\theta(1-\beta-\theta\lambda(\alpha+\beta))}{1+\theta\lambda}} = \chi \bar{A}_i^{\theta} \bar{u}_i^{\theta} y_i^{\frac{-\theta(1-\lambda)}{1+\theta\lambda}} l_i^{\frac{\theta(\alpha+1)}{1+\theta\lambda}} + \chi^{\frac{\theta\lambda}{1+\theta\lambda}} \sum_{j=1}^{N} \left(\bar{L}^{\lambda}\bar{t}_{ji}\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{A}_i^{\frac{\theta\lambda}{1+\theta\lambda}\theta} \bar{u}_i^{\theta} \bar{u}_j^{-\frac{\theta}{1+\theta\lambda}} y_j^{-\frac{\theta}{1+\theta\lambda}} l_j^{\frac{\theta(1-\beta)}{1+\theta\lambda}}. \tag{29}$$

An identical process for the urban model – starting from equilibrium conditions (19) and (20), substituting in equation (21) for the endogenous transportation costs, performing a matrix inversion, and incorporating endogenous traffic congestion from equation (24),(26), (27), – yields:

$$\left(l_i^R\right)^{1-\theta\beta} \left(l_i^F\right)^{\frac{\theta\lambda(1-\alpha\theta)}{1+\theta\lambda}} = \chi \bar{A}_i^{\theta} \bar{u}_i^{\theta} \left(l_i^F\right)^{\frac{\theta(\alpha+\lambda)}{1+\theta\lambda}} + \chi^{\frac{\theta\lambda}{1+\theta\lambda}} \sum_{j=1}^{N} \left(\bar{L}^{\lambda}\bar{t}_{ij}\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{u}_i^{\theta} \bar{A}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{u}_j^{-\frac{\theta}{1+\theta\lambda}} \left(l_j^R\right)^{\frac{1-\theta\beta}{1+\theta\lambda}} \tag{30}$$

$$\left(l_i^R\right)^{\frac{\theta\lambda(1-\beta\theta)}{1+\theta\lambda}} \left(l_i^F\right)^{1-\theta\alpha} = \chi \bar{u}_i^{\theta} \bar{A}_i^{\theta} \left(l_i^R\right)^{\frac{\theta(\beta+\lambda)}{1+\theta\lambda}} + \chi^{\frac{\theta\lambda}{1+\theta\lambda}} \sum_{j=1}^{N} \left(\bar{L}^{\lambda}\bar{t}_{ji}\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{A}_i^{\theta} \bar{u}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{A}_j^{-\frac{\theta}{1+\theta\lambda}} \left(l_j^F\right)^{\frac{1-\theta\alpha}{1+\theta\lambda}}. \tag{31}$$

Equations (28) and (29) for the economic geography model and equations (30) and (31) for the urban model determine the equilibrium distribution of economic activity $\{y_i, l_i\}$ or $\{l_i^F, l_i^R\}$ as a function of the model elasticities $\{\alpha, \beta, \theta, \lambda\}$, geography $\{\bar{A}_i, \bar{u}_i\}$, and fundamental infrastructure matrix $\bar{\mathbf{T}} \equiv [\bar{t}_{kl}]$, accounting for both the (standard) effect of transportation costs on the distribution of economic activity and the (new) effect of the distribution of economic activity on agents' optimal routing choice, the resulting traffic congestion, and the equilibrium transportation costs. Despite the complicated feedback loop between the two effects and the necessity of solving the resulting fixed point, the resulting equilibrium system is no more complicated than the typical system treating transportation costs as exogenous, as the number of equations and number of unknowns remains the same.

4.2 Existence and Uniqueness

While systems of equations with a structure as in (28) and (29) have, to our knowledge, not been studied previously, it turns out that the tools developed in Allen, Arkolakis, and Takahashi (2020) can be extended to analyze the properties of such an equilibrium.[17] We

[17] In the absence of traffic congestion, the equilibrium of a spatial equilibrium (e.g. equations (10) and (11) in the economic geography model) are examples of a system of non-linear integral equations known as a Hammerstein equation of the second kind, see e.g. Polyanin and Manzhirov (2008). Such systems, however,

first make an additional assumption on the infrastructure matrix:

Assumption 1. *The infrastructure matrix* $\bar{\mathbf{T}}$ *is strongly connected, i.e. there exists a path with finite costs between any two locations* i *and* j*, where* $i \neq j$*.*

Given Assumption 1, we summarize the results regarding existence and uniqueness in the following proposition:

Proposition 1. *For any strictly positive local geography* $\left\{\bar{A}_i > 0, \bar{u}_i > 0\right\}_{i \in \mathcal{N}}$*, aggregate labor endowment* $\bar{L} > 0$*, strongly connected infrastructure network* $\bar{\mathbf{T}} \equiv [\bar{t}_{kl}]$*, and model parameters* $\{\alpha \in \mathbb{R}, \beta \in \mathbb{R}, \theta > 0, \lambda \geq 0\}$*, then:*

1. *(Existence): There exists a strictly positive equilibrium.*
2. *(Uniqueness): For any* $\alpha \in [-1, 1]$ *and* $\beta \in [-1, 1]$*:*
   (a) *In an economic geography model with a symmetric infrastructure matrix, i.e.* $\bar{t}_{kl} = \bar{t}_{lk}$ *for all* $l \in \mathcal{N}$ *and* $k \in \mathcal{N}$*, the equilibrium is unique if:*
   $$\alpha + \beta \leq 0. \tag{32}$$
   (b) *In an urban model, the equilibrium is unique if:*
   $$\alpha \leq \frac{1}{2}\left(\frac{1}{\theta} - \lambda\right) \text{ and } \beta \leq \frac{1}{2}\left(\frac{1}{\theta} - \lambda\right) \tag{33}$$

Proof. See Appendix B. □

Despite the added complexity of endogenous traffic congestion, the sufficient conditions for uniqueness in the economic geography model are identical to those of an economic geography model with exogenous transportation costs, see Allen and Arkolakis (2014): productivity and amenity externalities must be (weakly) net dispersive to ensure a unique equilibrium. In the urban model, we achieve a similar result (here, however, because we do not impose symmetry, the productivity and amenity spillovers must be sufficiently dispersive individually, rather than combined). Unlike in the economic geography model, however, the strength of traffic congestion (λ) does play a role in ensuring uniqueness: the *stronger* the traffic congestion, the *more* dispersive the productivity and amenity externalities must be to ensure uniqueness. Unlike productivity and amenity externalities where the forces occur within a location, traffic congestion forces arise on flows between locations; loosely speaking, stronger traffic congestion forces can induce greater economic concentration by reducing the flows of goods or people between locations.

do not admit the inclusion of an endogenous additive term, as in (28) and (29) and also in (30) and (31).

4.3 Scale Dependence

In the absence of traffic congestion, equilibrium of the economic geography and urban models do not depend on the size of the aggregate labor endowment $\bar{L}$, i.e. both (standard) spatial models are *scale invariant.*[18] In the presence of traffic congestion, however, the equilibrium distribution of economic activity does depend on the size of the aggregate labor endowment $\bar{L}$, i.e. the equilibrium is *scale dependent*. As is evident from equations (28) and (29) (in the economic geography model) and equations (30) and (30) (in the urban model), increases in $\bar{L}$ are isomorphic to increases in costs of travel through the infrastructure network $\bar{t}_{ij}$, with an elasticity equal to the strength of the traffic congestion λ. Intuitively, the greater the aggregate labor endowment, the greater the traffic flowing through the network, and the greater the resulting traffic congestion. While the increases in the cost of travel through the infrastructure network are uniform, the impact on equilibrium transportation costs is not. To see this, we ask how a small uniform increase in the cost of travel through the entire infrastructure matrix by a factor of $c > 1$, i.e. suppose t_{kl} increases to ct_{kl}, changes equilibrium transportation costs (holding constant traffic congestion fixed). Differentiating equation (21) around $c = 1$ yields:

$$\frac{\partial \ln \tau_{ij}(c)}{\partial \ln c}|_{c=1} = \sum_{k=1}^{N}\sum_{l=1}^{N}\pi_{ij}^{kl},$$

i.e. a uniform increase in the cost of travel results in a non-uniform increase in bilateral transportation costs, where origins and destinations whose link intensity across the entire network is greater face the largest increases. These disproportionate changes in transportation costs alter the equilibrium distribution of economic activity, as the following example highlights.

4.4 Example

Consider a city comprising 25 locations arranged in a 5×5 grid, where, apart from their location in the grid, all locations are identical. Panel (a) of Figure 1 depicts the equilibrium distribution of economic activity in the absence of congestion forces (i.e. $\lambda = 0$). Locations in the center of the grid with better market access enjoy greater equilibrium economic activity (as indicated by taller "buildings"), and links in the center of the grid experience greater

[18] This fact is immediately evident from an examination of equations (10) and (11) (in the economic geography model) and equations (19) and (20) (in the urban model). In both systems, $\bar{L}$ only enters as a component of the endogenous scalar χ, so that any changes in $\bar{L}$ only changes $\bar{W}$ in such a way to ensure χ remains constant.

traffic (as indicated by their color), as they are more heavily used to travel through the network.

In panel (b), we introduce traffic congestion, setting $\lambda = 0.05$, but holding everything else constant. Traffic congestion disproportionately increases the cost of traversing the more heavily traveled central network segments. This disproportionately reduces the amount of traffic on those segments, causing relatively greater declines in central locations' market access and resulting in a fall in economic activity falls in the center of the city and rises in the outskirts: i.e. traffic congestion forces agents out of the center of the city and into the suburbs.

In panels (c) and (d), we increase the size of the economy from $\bar{L} = 100$ to $\bar{L} = 1000$ (panel c) and $\bar{L} = 10000$ (panel d). As discussed above, this would have no effect on the distribution of economic activity in the absence of traffic congestion, but in the presence of traffic congestion, scale matters. Increasing the aggregate population increases traffic everywhere, but the center of city is the worse affected: the resulting gridlock induces a reallocation of economic activity away from the center and toward the edges, further amplifying the move to the suburbs.

5 From Theory to Data

We now turn to applying our framework to evaluate the welfare impact of transportation infrastructure improvements. To do so, we begin by developing three helpful empirical tools: (1) we derive an equilibrium relationship between traffic flows on the one hand and trade (in the economic geography model) or commuting (in the urban model) on the other; (2) we show how to re-write the equilibrium conditions in terms of "exact hat" changes that depend only on observed traffic flows and economic activity and model parameters (e.g. the strength of traffic congestion); and (3) we present a procedure for estimating the strength of traffic congestion.

5.1 Traffic, Trade, and Commuting Flows

As we discussed in Section 3.2, there is a close link between the gravity equations for trade/commuting flows (equations 5 and 15, respectively) and the gravity equation for traffic (27). It turns out that this close link admits an analytical relationship between trade/commuting flows and traffic. Combining the two gravity equations (along with the definitions of the respective market access terms), one can express equilibrium trade flows in the economic geography model as:[19]

$$X_{ij} = c_{ij}^X \times Y_i \times E_j, \tag{34}$$

where c_{ij}^X is the $(i,j)^{th}$ element of the matrix $\mathbf{C}^X \equiv \left(\mathbf{D}^X - \boldsymbol{\Xi}\right)^{-1}$, $\mathbf{D}^X$ is a diagonal matrix with i^{th} element $d_i \equiv \frac{1}{2}\left(Y_i + E_i\right) + \frac{1}{2}\left(\sum_{j=1}^N \left(\Xi_{ji} + \Xi_{ij}\right)\right)$ and $\boldsymbol{\Xi} \equiv [\Xi_{ij}]$.

Similarly, one can express equilibrium commuting flows in the urban model as:

$$L_{ij} = c_{ij}^L \times L_i^R \times L_j^F, \tag{35}$$

where c_{ij}^L is the $(i,j)^{th}$ element of the matrix $\mathbf{C}^L \equiv \left(\mathbf{D}^L - \boldsymbol{\Xi}\right)^{-1}$, $\mathbf{D}^L$ is a diagonal matrix with i^{th} element $d_i \equiv \frac{1}{2}\left(L_i^R + L_i^F\right) + \frac{1}{2}\left(\sum_{j=1}^N \left(\Xi_{ji} + \Xi_{ij}\right)\right)$ and $\boldsymbol{\Xi} \equiv [\Xi_{ij}]$.

Equations (34) and (35) show that in both the economic geography and urban models, the equilibrium flows from origin to destination can be written only in terms of the economic activity in the origin (Y_i and L_i^R, respectively), economic activity in the destination (E_i and L_i^F, respectively), and the matrix of traffic flows through the network, $\boldsymbol{\Xi}$. In particular, equations (34) and (35), show that trade and commuting flows can be expressed as (an appropriately scaled) Leontief inverse of the traffic flows. Note that the expression depends only on available data and hence can be accomplished without knowledge of the underlying model elasticities. This result had two advantages, depending on the empirical availability of trade/commuting flows: in settings where both traffic flows and commuting / trade flows are observed (such as our empirical contexts below), it provides an out-of-sample test of the model, whereas in other settings where trade/commuting data are not available but traffic data is (e.g. much of the developing world), it still enables one to evaluate the welfare impacts of infrastructure improvements, a point we turn to next.

5.2 Counterfactuals

We now discuss how to evaluate the welfare impact of transportation infrastructure improvements in the presence of traffic congestion. Suppose we observe: (1) a matrix of traffic flows $\boldsymbol{\Xi} \equiv [\Xi_{ij}]$; (2) the distribution of economic activity in the origin and destination (i.e. (Y_i, E_j) in the economic geography model or $\left(L_i^R, L_j^F\right)$ in the urban model); and (3) the model parameters $\{\alpha, \beta, \theta, \lambda\}$. Suppose the observed economy has infrastructure network $\bar{\mathbf{T}} \equiv [\bar{t}_{kl}]$ and is in equilibrium. To determine how the distribution of economic activity and the welfare will change under an alternative infrastructure network $\bar{\mathbf{T}}' \equiv [\bar{t}'_{kl}]$, we follow the "exact hat algebra" approach pioneered by Dekle, Eaton, and Kortum (2008), where we denote with

[19] See Appendix A.3 for detailed derivations.

hats the change in variables, e.g. $\hat{\bar{t}}_{kl} \equiv \bar{t}'_{kl}/\bar{t}_{kl}$, $\hat{\gamma}_i \equiv \frac{\gamma'_i}{\gamma_i}$. For the economic geography model, one can write the equilibrium system of equations (10) and (11) in changes as:

$$\hat{y}_i^{\frac{1+\theta\lambda+\theta}{1+\theta\lambda}} \hat{l}_i^{-\frac{\theta(1+\alpha+\theta\lambda(\beta+\alpha))}{1+\theta\lambda}} = \hat{\chi}\left(\frac{E_i}{E_i+\sum_k \Xi_{ik}}\right)\hat{y}_i^{\frac{1+\theta\lambda+\theta}{1+\theta\lambda}}\hat{l}_i^{\frac{\theta(\beta-1)}{1+\theta\lambda}} + \hat{\chi}^{\frac{\theta\lambda}{1+\theta\lambda}}\sum_j\left(\frac{\Xi_{ij}}{E_i+\sum_k \Xi_{ik}}\right)\hat{\bar{t}}_{ij}^{-\frac{\theta}{1+\theta\lambda}}\hat{y}_j^{\frac{1+\theta}{1+\theta\lambda}}\hat{l}_j^{-\frac{\theta(1+\alpha)}{1+\theta\lambda}} \tag{36}$$

$$\hat{y}_i^{\frac{-\theta(1-\lambda)}{1+\theta\lambda}} \hat{l}_i^{\frac{\theta(1-\beta-\theta\lambda(\alpha+\beta))}{1+\theta\lambda}} = \hat{\chi}\left(\frac{Y_i}{Y_i+\sum_k \Xi_{ki}}\right)\hat{y}_i^{\frac{-\theta(1-\lambda)}{1+\theta\lambda}}\hat{l}_i^{\frac{\theta(\alpha+1)}{1+\theta\lambda}} + \hat{\chi}^{\frac{\theta\lambda}{1+\theta\lambda}}\sum_{j=1}^N\left(\frac{\Xi_{ji}}{Y_i+\sum_k \Xi_{ki}}\right)\hat{\bar{t}}_{ji}^{-\frac{\theta}{1+\theta\lambda}}\hat{y}_j^{-\frac{\theta}{1+\theta\lambda}}\hat{l}_j^{\frac{\theta(1-\beta)}{1+\theta\lambda}} \tag{37}$$

Similarly, for the urban model, the equilibrium system defined by equations (19) and (20) can be written in changes as:

$$\left(\hat{l}_i^R\right)^{1-\theta\beta}\left(\hat{l}_i^F\right)^{\frac{\theta\lambda(1-\alpha\theta)}{1+\theta\lambda}} = \hat{\chi}\left(\frac{L_i^F}{L_i^F+\sum_k \Xi_{ik}}\right)\left(\hat{l}_i^F\right)^{\frac{\theta(\alpha+\lambda)}{1+\theta\lambda}} + \hat{\chi}^{\frac{\theta\lambda}{1+\theta\lambda}}\sum_{j=1}^N\left(\frac{\Xi_{ij}}{L_i^F+\sum_k \Xi_{ik}}\right)\hat{\bar{t}}_{ij}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_j^R\right)^{\frac{1-\theta\beta}{1+\theta\lambda}} \tag{38}$$

$$\left(\hat{l}_i^R\right)^{\frac{\theta\lambda(1-\beta\theta)}{1+\theta\lambda}}\left(\hat{l}_i^F\right)^{1-\theta\alpha} = \hat{\chi}\left(\frac{L_i^R}{L_i^R+\sum_k \Xi_{ki}}\right)\left(\hat{l}_i^R\right)^{\frac{\theta(\beta+\lambda)}{1+\theta\lambda}} + \hat{\chi}^{\frac{\theta\lambda}{1+\theta\lambda}}\sum_{j=1}^N\left(\frac{\Xi_{ji}}{L_i^R+\sum_k \Xi_{ki}}\right)\hat{\bar{t}}_{ji}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_j^F\right)^{\frac{1-\theta\alpha}{1+\theta\lambda}} \tag{39}$$

Both systems of equilibrium changes bear a close resemblance to their level variants above; the key distinction, however, is that the local geography and the infrastructure matrix are replaced with shares that depend only on the observed traffic flows and observed distribution of economic activity. Given such data – and knowledge of the model parameters $\{\theta, \alpha, \beta, \lambda\}$ – it is then straightforward to evaluate the impact of any transportation infrastructure improvements $\left\{\hat{\bar{t}}_{ij}\right\}$ on the equilibrium distribution of economic activity and aggregate welfare.[20] While the first three parameters $\{\theta, \alpha, \beta\}$ are familiar ingredients in spatial models (and we will be calibrating their values to those of the literature below), the strength of traffic congestion λ is new to our framework. We turn now to its estimation.

5.3 Estimating the Strength of Traffic Congestion

To derive a straightforward estimating equation for the strength of the endogenous traffic congestion, we make two additional assumptions. First, we follow an extensive literature on trade cost estimation, and assume that transportation costs t_{kl} are a log-linear function of travel time.[21] As a result, we can write t_{kl} as a function of the distance of the link and the

[20] Online Appendix E describes the algorithm used to solve equations (19) and (20) given these ingredients.

[21] For example, Hummels and Schaur (2013) find that time is an important component of international trade costs, and Pascali (2017) and Feyrer (2019) use plausibly exogenous shocks to travel time as instruments

speed of travel on the link:

$$t_{kl} = \left(distance_{kl} \times speed_{kl}^{-1}\right)^{\delta_0}, \tag{40}$$

where δ_0 is the time elasticity of the transportation cost. In our preferred results below, we set $\delta_0 = 1/\theta$ to imply a "distance elasticity" of negative one, which is consistent with a large gravity literature, see e.g. Disdier and Head (2008) and Chaney (2018).[22]

Our second assumption is that time per unit distance (inverse speed) is a log-linear function of traffic congestion (measured as total vehicle miles traveled per lane-miles, or equivalently, traffic per average lanes) as follows:

$$speed_{kl}^{-1} = m_0 \times \left(\frac{\Xi_{kl}}{lanes_{kl}}\right)^{\delta_1} \times \varepsilon_{kl} \tag{41}$$

where $\delta_1$is the congestion elasticity of inverse speed, m_0 is the average rate of flow without congestion, $lanes_{kl}$ are the average number of lanes on a link, and ε_{kl} is a segment specific idiosyncratic free rate of flow. The log-linear specification was first posited by Vickrey (1967), and while simple, has a number of advantages in our setting.[23] First, combined with equations (40), and (41) immediately implies:

$$t_{kl} = \bar{t}_{kl} \times (\Xi_{kl})^{\lambda},$$

where $\bar{t}_{kl} \equiv lanes_{kl}^{-\delta_0\delta_1} \times (distance_{kl} \times m_0 \times \varepsilon_{kl})^{\delta_0}$ and $\lambda \equiv \delta_0\delta_1$. That is, this simple setup offers a micro-foundation for the traffic congestion formulation (25) posited in Section 4. Second, treating distance and the free rate of flow as segment specific time-invariant characteristics, equation (41) provides a simple relationship between infrastructure improvements and the change in the infrastructure matrix:

$$\hat{\bar{t}}_{kl} = \hat{lanes}_{kl}^{-\lambda}. \tag{42}$$

for changes in trade costs. Anderson and Van Wincoop (2004) note that the assumption that trade costs are a log-linear function of distance – a special case of our assumption when speed of travel is constant – is "by far the most common assumption" (p.710).

[22]In Online Appendix G, we present alternative results where we estimate δ_0 by using the estimated distance elasticity from gravity equations of our observed trade and commuting flows, respectively, on distance, which imply slightly stronger traffic congestion forces (as our estimates of the distance elasticity are 1.6 in the economic geography case and 1.45 in the urban case). As is evident, the welfare impacts of infrastructure improvements are qualitatively and quantitatively similar to those with our preferred estimates.

[23]Vickrey (1967) assumes a log-linear relationship between inverse speed and traffic congestion, where inverse speed is defined relative to an unimpeded inverse speed (see his equation 1). In equation (41) there is no such unimpeded inverse speed, i.e. while we follow Vickrey (1967) in considering a log-linear approximation of the impact of congestion on travel time, our approximations centers around an inverse speed of zero rather than the free-flow rate of travel.

As additional lane-miles are added to a segment, congestion on the segment falls, reducing the exogenous component of transportation costs with an elasticity of λ. This is intuitive: the greater the strength of traffic congestion, the larger the impact of adding additional lanes. However, it is important to (re-)emphasize that improvements in the infrastructure matrix will also result in an endogenous increase in traffic demand. Indeed, combining equation (42) with (27), we see that the elasticity of traffic to lanes is $\frac{\partial \ln \Xi_{kl}}{\partial \ln lanes_{kl}} = \frac{\lambda\theta}{1+\lambda\theta}$, i.e. the limiting case as traffic congestion becomes infinitely large is that traffic increases proportionately with the adding of additional lanes, as in "the fundamental law of road congestion" identified by Duranton and Turner (2011).

The final advantage of this setup is that it delivers a straightforward estimating equation and, combined with the traffic gravity equation (27), an appropriate identification strategy. Taking logs of equation (41) yields:

$$\ln speed_{kl}^{-1} = \ln m_0 + \delta_1 \ln\left(\frac{\Xi_{kl}}{lanes_{kl}}\right) + \ln \varepsilon_{kl}, \tag{43}$$

i.e. a regression of inverse speed on traffic congestion can in principal identify the congestion elasticity of inverse speed δ_1. An ordinary least squares regression is inappropriate in this case, as the residual is the free rate of flow on the segment kl, which enters into $\bar{t}_{kl}$ and so, from the traffic gravity equation (27) is negatively correlated with traffic Ξ_{kl}, biasing the estimate of δ_1 downwards. Instead, we propose to use an instrumental variables strategy, instrumenting for traffic Ξ_{kl} with observables that affect traffic demand for a segment but are uncorrelated with the free rate of flow on the segment. From the traffic gravity equation (27), conditional on k and l fixed effects, any component of $\bar{t}_{kl}$ that does not affect the free rate of flow is a suitable instrument. Intuitively, we can use observables that shift the traffic gravity (demand) equation to identify the slope of the traffic congestion (supply) equation. We describe such instruments in the next section, where we apply our procedure to determine the welfare impact of transportation infrastructure improvements in two different settings.

6 The welfare impact of transportation infrastructure improvements

We first apply the economic geography variant of our framework to evaluate the welfare impact (and, given cost estimates, the return on investment) of small improvements to every single segment of the U.S. Interstate Highway network. We then apply the urban variant of our framework to do the same for each segment of the road network in Seattle, WA.[24]

6.1 Traffic across the Country: The U.S. Highway Network

The U.S. National Highway System is largest highway system in the world. The main backbone of the National Highway System – the Interstate Highway System – is one of the world's largest infrastructure megaprojects in history (Kaszynski, 2000), taking more than thirty five years to construct at an estimated cost $650 billion (in 2014 dollars), and total annual maintenance costs are approximately $70 billion (CBO, 1982; FHA, 2008; NSTIFC, 2009; ASCE, 2017). However, little is known about the relative importance of different segments of the highway system in terms of how each affects the welfare of the U.S. population. Such knowledge is crucial for appropriately targeting future infrastructure investments.

Our strategy to estimate the welfare impact of improvements to the U.S. Highway System is straightforward: for each segments of the network, we will use equations (36) and (37) for the economic geography variant of our approach to estimate the aggregate welfare impact ($\hat{\bar{W}} = \hat{\chi}^{-\frac{1}{\theta}}$) of a small (1%) improvement to the infrastructure network. We then use equation (42) to calculate how many lane-miles must be added in order to achieve a 1% improvement in order to estimate such an infrastructure cost. Given costs and benefits, we can then identify the highway segments with the greatest return on investment. This procedure requires just two ingredients: (1) data on traffic $\{\Xi_{kl}\}$ and income $\{Y_i = E_i\}$; and (2) knowledge of the four model parameters $\{\theta, \alpha, \beta, \lambda\}$. We discuss the source of these ingredients in turn.

6.1.1 Data

We briefly summarize the data used here; see Online Appendix F.1 for more details. The primary source of data we use to construct the infrastructure network is the 2012 Highway Performance Monitoring System (HPMS) dataset by the Federal Highway Administration. This dataset comprises the length, location, number of lanes, and average annual daily traffic (AADT) over 330,021 segments of the U.S. highway system.[25]

To create the infrastructure network, we begin by placing nodes at each endpoint and intersection between two different Interstate highways and collapsing all nodes within the same core-based statistical area (CBSA) to a single CBSA point. This resulting 228 locations

[24] We leave the evaluation of large scale changes to the infrastructure network – while feasible using the methodology presented in 6 – to future work.

[25] The traffic data is reported for a segment without reference to the direction of travel. Combined with the fact that we impose $Y_i = E_i$ in the data, this results in two implications: first, as equations (36) and (37) have symmetric kernels, the uniqueness results of Proposition 2(a) apply to the counterfactuals conducted; second, to be consistent with the data, we examine infrastructure improvements that symmetrically improve a segment in both directions of travel.

and 704 links between adjacent nodes, where for each link we construct a length-weighted average of AADT and number of lanes. Panel (a) of Figure 3 depicts the actual highway network and the resulting infrastructure network.

To this network, we append four additional data sources. First, to estimate the strength of congestion, we recover the time of travel ($time_{kl}$) across each link from the HERE API using the georoute Stata command by Weber and Péclat (2017). Second, we calculate the population and income at each node by summing the population and averaging the median income of all cities from Edwards (2017) (which is itself based on the U.S. Census and American Community Survey) within 25 miles of the node. Third, we estimate the cost of improving each link based on the topography of its constituent segments. To do so, we classify each segment of the Interstate Highway System into one of seven categories from the Federal Highway Administration's Highway Economic Requirements System (HERS) Federal Highway Administration (2015), each of which is associated with an estimated cost of adding one lane-mile.[26] To determine the average cost of adding one lane-mile to a link, we construct a distance-weighted average of the cost of improving each of its constituent segments. Fourth, we rely on the 2012 Commodity Flow Survey (CFS) to construct measures of the value of bilateral trade flows between each CBSA; for CFS areas comprising more than one CBSA, we allocate observed CFS area flows to CBSAs proportionally to their share of the CFS area's total income.

6.1.2 Predicted versus observed trade flows

As a first check of the validity of the framework developed above, we compare the observed value of bilateral trade flows between CBSAs from the CFS to the predicted bilateral trade flows arising from observed traffic flows using equation (34). To do so, we assume that each element of the matrix of traffic flows $\boldsymbol{\Xi} \equiv [\Xi_{kl}]$ is equal to the observed AADT along the highway segment, which is equivalent to assuming that each car is carrying a value of trade equal to the average value of a single individual's labor. This of course abstracts from many nuances of traffic flows, including shipments via truck (where the trade value exceeds this average) as well as traffic for non-trade purposes such as commuting and shopping (where the trade value falls below this average). Given these abstractions, it is all the more remarkable how well traffic across the interstates is able to predict actual trade between CBSAs. Panel

[26] The Federal Highway Administration provides seven different cost categories for the interstate highway system that we can use based on geographical characteristics and urbanization: rural-flat (\$1.923m), rural-rolling (\$2.085m), rural-mountainous (\$6.492m), small-urban (\$3.061m), small-urbanized (\$3.345m), large-urbanized (\$5.598m), major-urbanized (\$11.197m). We are grateful to the experts at the U.S. Department of Transportation Volpe National Transportation Systems Center for their substantial assistance in developing these cost estimates.

(a) of Figure 2 shows the scatter plot between observed and predicted (log) trade flows, conditional on origin and destination fixed effects (so the only variation arises from the bilateral flows and not e.g. income in the origin or destination). As is evident, there is a strong positive correlation of 0.60, indicating the traffic matrix – through the lens of the theory and despite obvious measurement issues – does a good job of predicting trade flows.

6.1.3 Estimation

We now discuss our choice of the four model parameters $\{\theta, \alpha, \beta, \lambda\}$. As the first three model parameters – the trade elasticity θ, productivity externality α, and amenity externality β – are standard in the economic geography literature, we choose central values from the literature. We set $\theta = 8$ to match previous estimates of the trade elasticity.[27] We also choose $\alpha = 0.1$, and $\beta = -0.3$, which corresponds to the estimated scale economies found in the literature, as e.g. summarized in Rosenthal and Strange (2004) and Combes and Gobillon (2015) and the share of consumption allocated to housing, see e.g. Allen and Arkolakis (2014).[28] From Proposition 1, this choice of parameter values guarantees the existence of a unique equilibrium.

To estimate the strength of traffic congestion, we follow the estimation procedure described in Section 5.3, regressing observed inverse speed on (appropriately instrumented) traffic congestion as in equation (43). As implied by the traffic gravity equation (27), recall that an appropriate instrument would be something that – conditional on start-location and end-location fixed effects – affects the cost of travel $\bar{t}_{kl}$ but is uncorrelated with the free-flow speed of travel on the link. In the context of the U.S. highway system, we propose that the distance along the link is such an appropriate instance. Distance clearly affects the cost of travel (and so is relevant), and given the relative homogeneity of U.S. highways in terms of speed limits, lanes, limited access, etc., we have no reason to believe that longer or shorter links have different free flow rates of speed (so it is plausibly excludable).

Panel A of Table 1 presents the results. Columns (1) and (2) show using OLS that there is a positive, but small, correlation between inverse speed of travel and congestion. Column (3) presents the first stage regression of traffic on distance: as expected, conditional on

[27] Donaldson and Hornbeck (2016) estimate a trade elasticity of 8.22 for U.S. intra-national trade, albeit in the late 19th century. Eaton and Kortum (2002) estimate a trade elasticity between 3.60 and 12.86 for international trade, with a preferred estimate of 8.28.

[28] In reviews of the literature, Rosenthal and Strange (2004) and Combes and Gobillon (2015) conclude that agglomeration elasticities at the city level are likely between 0.03 and 0.08. As in Allen and Arkolakis (2014), we choose a spillover of $\alpha = 0.1$ to also incorporate the effects of entry on overall output. As robustness, in Online Appendix G, we repeat the exercise for alternative constellations of these model parameters, including (1) removing the externalities, (2) lowering the trade elasticity; and (3) increasing the traffic congestion parameters. As is evident, both the patterns of welfare elasticities and the returns on investment are both qualitatively and quantitatively similar to the results presented here.

start-location and end-location fixed effects, distance is strongly negatively correlated with traffic. Column (4) presents the IV regression: Consistent with OLS exhibiting downward bias due to traffic demand being lower on slower links, the IV is substantially larger, finding a coefficient $\delta_1 = .739$ (with standard error of .181). Recall from above that we set $\delta_0 = 1/\theta$ to match the unit distance elasticity, so this implies $\lambda = \delta_1\delta_0 = 0.092$, i.e. a 10% increase in traffic flows is associated with a 7.4% increase in travel time, resulting in a 0.9% increase in the transportation cost.

6.1.4 Results

Given the observed traffic data and estimated parameters, we calculate the aggregate welfare elasticity to a 1% reduction in iceberg transportation costs on every link (in both directions of travel) of the U.S. Highway System using equations (36) and (37), i.e. $\frac{1}{2}\left(\frac{\partial \ln \bar{W}}{\partial \ln \bar{t}_{kl}} + \frac{\partial \ln \bar{W}}{\partial \ln \bar{t}_{lk}}\right)$. Panel (a) of Figure 4 presents our results. While all highway segments have positive welfare elasticities, the elasticities are largest on short segments connecting CBSAs in densely populated areas, e.g. along I-95 between Boston and Philadelphia and on I-5 between Los Angeles and San Diego. Welfare elasticities are also large along longer highway segments that do not directly connect large urban areas but that are major thoroughfares for trade, e.g. in the interstates passing through Indiana ("the crossroads of America"). Conversely, highway segments that neither connect major urban areas nor are used intensively for trade – such as I-90 through Montana – have the lowest positive impact on aggregate welfare.

How much does incorporating endogenous traffic congestion affect our welfare elasticity estimates? Panel (a) of Figure 5 presents a scatter plot of the welfare elasticity for each segment with and without congestion. Not surprisingly, in the absence of traffic congestion, the welfare gains from reducing transportation costs are greater. What is surprising, however, is that there is substantial variation in welfare gains with and without congestion across segments, highlighting that traffic congestion plays an important role in determining which segments would achieve the greatest welfare gains.

The benefit of improving a link, of course, is only half of the story. To calculate an return on investment, we pursue a cost-benefit approach. On the benefit side, we translate the welfare elasticity into a dollar amount use a compensating variation approach, asking how much the annual U.S. real GDP (of \$19 trillion) would have to increase (in millions of chained 2012 US dollars) to bring about the same welfare increase we estimate. On the cost side, we first use equation (42) to calculate how many additional lane-miles would need to be added to the route to achieve a 1% reduction in transportation costs. We then take multiple this number of lane-miles by the cost per lane-mile to get a total construction cost. We assume a 20 year depreciation schedule (as in Appendix C of Office of the State Auditor

(2002)), a 5% annual maintenance cost, and a 3% borrowing cost, which together imply 10% of the construction cost is incurred each year.[29]

Panel (a) of Figure 6 reports the annual return on investment (RoI) for each segment of the U.S. highway system. On average, infrastructure improvement return are well-worth the investment, with a mean RoI of just over 108%. However, there is also huge variance in returns, with some segments offering negative RoI (such as I-90 through Montana) and others offering much higher than average. Panel (a) of Table 2 presents the ten links with the highest RoI (each of which exceed 400%). All ten are for links outside the largest cities, where reducing transportation costs is less costly. This does not mean that returns are entirely driven by costs: the links with the highest returns are those on the periphery of densely populated areas with high welfare elasticities, reflecting the importance of trade between these regions.

6.2 Traffic in the City: The Seattle Road Network

We now provide use the urban variant of our framework to examine the welfare impacts of transportation infrastructure improvement in Seattle, WA. Seattle provides an ideal test-case for our framework for several reasons, notably: (1) it has some of the worst traffic in the U.S.; (2) with limited (non-bus) public transit options, its road network plays a critical role in commuting; and (3) its road network is particularly interesting, with multiple natural choke points created by the waterways which intersect the city.[30]

Our strategy for estimating the welfare impacts of improvements to the Seattle road network proceeds analogously to the U.S. highway system above: for each link in the road network, we estimate the change in the aggregate welfare $\left(\hat{\bar{W}} = \hat{\chi}^{-\frac{1}{\theta}}\right)$ from a small (1%) improvement using equations (38) and (39). Doing so requires just two ingredients: (1) data on traffic (Ξ_{kl}), residential population $\left(L_i^R\right)$, and workplace population $\left(L_i^F\right)$ and (2) values for the model parameters $\{\theta, \alpha, \beta, \lambda\}$. We discuss the source of both ingredients in turn.

[29] Annual spending equal to 10% of total cost accords well with various sources. Feigenbaum, Fields, and Purnell (2020) find the average total-disbursements of state-controlled highway in 2018 is \$308,558 per lane-mile, 8.5% of our length-weighted average estimated construction cost of \$3.6m per lane-mile. ASCE (2017) find in 2014 that states spent \$70 billion in maintenance and upkeep of the highway system, 10.7% of the \$650 billion construction cost of the interstate highway system.

[30] A 2019 study by Apartment Guide ranked Seattle as the second worst city for commuters; a coauthor vividly remembers running out of gas while stuck in Seattle traffic. Of commuters, over half drive alone or carpool. Of those that use public transit, the vast majority of trips are conducted via buses: Commute Seattle's 2016 Center City Commuter Mode Split Survey found that, among public transit commuters, over three-quarters take the bus while only around a sixth take the train (or light rail or streetcar) (EMC Research, 2016).

6.2.1 Data

We briefly summarize the data used here; see Online Appendix F.2 for more details. Data on the location, functional system (i.e., interstate, arterial road, local road, etc.), ownership, AADT, lane width, and possibility for lane expansion of the 9,188 road segments within the municipal boundaries of Seattle were taken from the 2016 HPMS release for the state of Washington.[31] To construct our adjacency matrix of Seattle, we divide Seattle into ˜1 sq. mi. grids,[32] place the center point of each of these grids as a node into ArcGIS Network Analyst, and find the least-cost path between each of these nodes. This gives us a total of 217 nodes, with 1,384 links between adjacent nodes, 1,338 for which we observe traffic.[33] Panel (b) of Figure 3 depicts the actual Seattle road network and the resulting infrastructure network.

We append to this network five additional sources of data. First, we calculate the time of travel between each link from the HERE API using the georoute Stata command by Weber and Péclat (2017). Second, we observe the labor force and residential population density at the census block group level from the 2017 Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES), which we aggregate to our constructed grids (allocating population from block groups intersected by our grids proportional to the area of the block group within each grid). Third, the LODES data also provide bilateral commuting flows between census block groups, which we aggregate to bilateral grid cell pairs using a similar procedure. Fourth, we estimate the cost of adding an additional lane-mile to each link in the network. To do so, we classify each Seattle's road sections into the major urbanized road type based on the population of the Seattle urban area (as defined by the Census Bureau's 2012 Urban Area data) and additionally indicate if the section is "restricted" if the HPMS indicate that additional lanes cannot be added. Then, based on a road section's functional system classification, its major urbanized classification, and whether it is a high cost road to improve or not, we code each road section with the cost of adding a lane-mile to

[31] Traffic data on a road segment is reported without regard to the direction of travel. As such, we evaluate simultaneous improvements to each link in the Seattle road network in both directions of travel. This has the added advantage of reconciling our urban framework – where traffic is modeled as flowing from an agents' residence to her workplace – to the (presumed) empirical reality that the agent returns home after work.

[32] This approach is necessary because, at this level, typical units of observation like census blocks and block grounds are endogenous to the road structure of Seattle; this leaves us with concerns that census blocks which are larger are in a less dense area of Seattle with less traffic.

[33] Unlike the interstates, where we observe all segments of the highway system, our analysis does not cover every road in Seattle, just those along the least-cost path between adjacent nodes. We do, however, observe the entirety of the Seattle road network in our dataset. We assume the route along the least-cost path between nodes reasonably captures the costs of moving across similar paths, on different roads, between the same nodes.

it, as estimated by the FHA's HERS from Federal Highway Administration (2015).[34] Fifth, we calculate the number of intersections and turns along each link of the network using the ArcGIS network analyst.

6.2.2 Predicted versus observed commuting flows

As a first pass of the validity of the urban variant of our framework to the data, we compare the observed bilateral commuting flows from LODES to those predicted from the observed traffic flows using equation (35). To do so, we assume that each element of the matrix of traffic flows $\Xi \equiv [\Xi_{kl}]$ is equal to the observed AADT along that road segment. This assumes every vehicle carries one commuter. As with the interstates, this introduces obvious measurement error: some vehicles contain many commuters (e.g. buses), whereas other vehicles contain none (e.g. when driving to go shopping). And like with interstates, it is remarkable how well observed traffic flows are able to predict commuting flows, as panel (b) of Figure 2 illustrates. Even conditional on origin and destination fixed effects, there is a positive correlation between predicted and observed commuting flows of 0.43, indicating that the urban model with traffic congestion is able to successfully predict observed commuting flows.

6.2.3 Estimation

We now discuss our choice of model parameters $\{\theta, \alpha, \beta, \lambda\}$. As the first three model parameters are standard in the quantitative urban literature, for our preferred estimates presented here we set them equal to the values of estimated in the seminal work of Ahlfeldt, Redding, Sturm, and Wolf (2015), with $\theta = 6.83$, $\alpha = -0.12$, and $\beta = -0.1$.[35] From Proposition 1, this choice of parameter values guarantees the existence of a unique equilibrium.

To estimate the strength of traffic congestion, we again proceed as discussed in Section (39), regressing the observed inverse speed of travel over a link on the traffic congestion,

[34] For major urban areas, the Federal Highway Administration provides the following estimates of the cost of adding an additional lane-mile: for interstates/freeways (\$11.197m when unrestricted, \$46.691m when restricted), other principal arterial (\$8.252m when unrestricted, \$31.988m when restricted), and minor arterial/collector (\$5.614m when unrestricted and \$31.988m when restricted. Further details are in Online Appendix F.2.1.

[35] Ahlfeldt, Redding, Sturm, and Wolf (2015) also allow for externalities to affect nearby locations, which they estimate to steeply decay over space; here we assume externalities have only local effects. Our choice of $\alpha = -0.12$ combines their estimated agglomeration externality with the congestion force that arises from floor space being used in the production of goods. As robustness, in Online Appendix G, we repeat the exercise for alternative constellations of these model parameters where we vary the commuting elasticity, strength of externalities, and strength of congestion. As with the analysis of the U.S. highway network, both the patterns of welfare elasticities and the returns on investment are both qualitatively and quantitatively similar to the results presented here.

appropriately instrumented by a demand shifter uncorrelated with the free-flow rate of speed over the link. Unfortunately, the instrument used for the U.S. highway system – distance – is inappropriate in a city setting. There exists enormous variation in the types of roads and speed of travel within Seattle (e.g. surface streets with stop signs, larger streets with major intersections, highways, etc.), so it is likely that the distance of a segment is correlated with its free-flow rate of speed (e.g. a link which travels along a highway might be longer but faster). As an alternative, we propose that the complexity of a route is a suitable instrument: conditional on the free-flow rate of speed, drivers would prefer to take routes that are less complex. To measure complexity, we use the number of turns along the route as our instrument, conditioning on the number of intersections.[36] Intuitively, intersections reduce the free-flow rate of speed of travel regardless if one turns or not, while turns themselves present an additional inconvenience to drivers.

Panel (b) of Table 1 presents the results. Column (1) shows that there is actually a small negative correlation between inverse speed and traffic, consistent with substantial downward bias due to the heterogeneity in free-flow speed across links (e.g. faster links on highways also have higher traffic). Column (2) presents the first stage results; as expected, the greater the number of turns along a route (conditional on the number of intersections), the lower the traffic along that link. Column (3) presents the IV results, where we estimate $\delta_1 = 0.118$ (with a standard error of 0.048). One potential concern with the instrument is that controlling for the number of intersections alone may not be sufficient to allay the concern that more complex routes are more likely to travel over smaller (and slower) roads. In Columns (4) and (5) present the first and second stage results where we nonparametrically control for the share of the route that travels over arterial and local roads.[37] Such a procedure compare links with similar road compositions, mitigating the concern that route complexity is correlated with unobserved speed of travel. Adding these controls increases our estimate of $\delta_1 = 0.488$ (with standard error of 0.278). Combined with the maintained assumption that $\delta_0 = 1/\theta$ (to generate a unit distance elasticity), this implies a traffic congestion parameter of $\lambda = \delta_1 \delta_0 = 0.071$, i.e. a 10% increase in traffic flows is associated with a 4.9% increase in travel time, resulting in a 0.7% increase in the transportation cost. It is interesting to note that while the elasticity of travel time to congestion is smaller in Seattle than U.S. highways – perhaps due to the lower free-flow rates of speed within a city – the impact of traffic congestion on transportation costs in both settings is quite similar.

[36]See Online Appendix Figure F.1 for an example of how the instrument is constructed.

[37]To do so, we include fixed effects for each decile of arterial and local road shares.

6.2.4 Results

For each link in the road network, we simulate a 1% reduction in transportation costs in each direction and calculate the change in aggregate welfare elasticity $\frac{1}{2}\left(\frac{\partial \ln \bar{W}}{\partial \ln \bar{t}_{kl}} + \frac{\partial \ln \bar{W}}{\partial \ln \bar{t}_{lk}}\right)$. Panel (b) of Figure 4 presents our findings. While a reduction in transportation costs on all links are welfare improving, the largest welfare elasticities are greatest in the center of the city (downtown). Welfare elasticities are also higher for the various choke-points in the road network (oftentimes corresponding to bridges over water).

Panel (b) of Figure 4 compares these estimated welfare elasticities to those estimated without traffic congestion. As with the U.S. highway system, ignoring congestion would not result in overestimates of the welfare elasticities, it would also substantially change which links one would identify as having the largest welfare effects. For example, the link at the top left of the figure corresponds to a highly trafficked stretch of interstate I-5. Ignoring traffic congestion would cause one to identify this stretch as the one whose improvement would yield the greatest welfare gains for the city. Accounting for the endogenous change in traffic congestion throughout the whole network, the aggregate welfare elasticity to improving this link is not even in the top fifty of links.

Finally, we combine these welfare elasticities with estimated costs of construction to estimate a return on investment for each link of the Seattle road network. We proceed analogously to the U.S. highway system case, first calculating the necessary lane-miles to achieve a 1% reduction in transportation costs, assuming 10% of construction costs are incurred each year, and then using a compensating variation approach to assign a dollar value to the aggregate welfare gains.[38] We find that improving the average link in Seattle yields an annual return of 16.8% for the residents of the city. Like with the U.S. highway system, however, there is substantial heterogeneity, with returns varying from less than 25% to more than 250%. Panel (b) of Figure (6) shows the RoI for each segment; the highest returns are concentrated in the center of the city. Panel (b) of Table 2 lists the top 20 links in terms of their RoI; half of the list are either entirely within downtown Seattle or between downtown Seattle and another part of the city. Other locations with high returns on infrastructure improvement include the area around the University of Washington campus and Lake City Way in the neighborhood of North Seattle. On the other hand, we estimate that nearly half (331 of 692) links in the Seattle road network would generate *negative* returns of investment, highlighting the importance of well-targeted infrastructure improvements.

[38] To identify a "GDP" for the municipality of Seattle, we sum over the incomes of all our grid cells, which we derive from block group income measures from the American Community Survey. We estimate a GDP of $45.5 billion.

7 Conclusion

This paper proposes a new spatial framework that incorporates traffic congestion and uses it to evaluate the welfare impact of transportation infrastructure improvements. In doing so, it combines the rich geography and general equilibrium structure of existing quantitative spatial models with the endogenous routing and traffic congestion of transportation models, but where both the distribution of economic activity and the resulting traffic patterns are determined jointly in equilibrium.

The approach generates analytical expressions for transportation costs between any two locations, the traffic along each link of the transportation network, and the equilibrium distribution of economic activity across the economy. This tractability not only allows us to characterize the equilibrium properties of the framework, but it also facilitates applying the framework to evaluate the welfare impacts of transportation infrastructure improvements empirically. Using readily available traffic data we show that for both the U.S. highway network and the Seattle road network, where you improve the road network matters, as there are large differences in returns on investment across different links.

The goal of this paper has been to provide a tractable framework that bridges the gap between the quantitative spatial and transportation economics literatures. As a result, we hope it can facilitate the answering of a number interesting and unresolved research questions, including: How does traffic congestion impact urban land use? What is the best way to design congestion tolls? How does the presence of multiple uses of transportation infrastructure (e.g. trade, commuting, consumption) interact in determining traffic congestion and the spatial distribution of economic activity? We look forward to fruitful future research on these topics.

References

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

Akamatsu, T. (1996): "Cyclic flows, Markov process and stochastic traffic assignment," *Transportation Research Part B: Methodological*, 30(5), 369–386.

Allen, T., and C. Arkolakis (2014): "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*.

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

American Association of State Highway and Transportation Officials (2016): "A Policy on Design Standards - Interstate System," Discussion paper.

Anderson, J. E., and E. Van Wincoop (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93(1), 170–192.

——— (2004): "Trade Costs," *Journal of Economic Literature*, 42(3), 691–751.

ASCE (2017): "2017 infrastructure report card," .

Beckmann, M. J., C. B. McGuire, and C. B. Winsten (1955): "Studies in the Economics of Transportation," .

Bell, M. G. (1995): "Alternatives to Dial's logit assignment algorithm," *Transportation Research Part B: Methodological*, 29(4), 287–295.

CBO (1982): "The interstate highway system: Issues and options," *Congressional Budget Office of the United States*.

Chaney, T. (2018): "The gravity equation in international trade: An explanation," *Journal of Political Economy*, 126(1), 150–177.

Chartrand, G. (1977): *Introductory graph theory*. Courier Corporation.

Combes, P.-P., and L. Gobillon (2015): "The empirics of agglomeration economies," in *Handbook of regional and urban economics*, vol. 5, pp. 247–348. Elsevier.

Costinot, A., and A. Rodríguez-Clare (2014): "Trade theory with numbers: Quantifying the consequences of globalization," in *Handbook of international economics*, vol. 4, pp. 197–261. Elsevier.

De Palma, A., M. Kilani, and R. Lindsey (2005): "Congestion pricing on a road network: A study using the dynamic equilibrium simulator METROPOLIS," *Transportation Research Part A: Policy and Practice*, 39(7-9), 588–611.

De Palma, A., R. Lindsey, E. Quinet, and R. Vickerman (2011): *A handbook of transport economics*. Edward Elgar Publishing.

Dekle, R., J. Eaton, and S. Kortum (2008): "Global Rebalancing with Gravity: Measuring the Burden of Adjustment," *IMF Staff Papers*, 55(3), 511–540.

Disdier, A.-C., and K. Head (2008): "The puzzling persistence of the distance effect on bilateral trade," *The Review of Economics and statistics*, 90(1), 37–48.

Donaldson, D. (2015): "The gains from market integration," *economics*, 7(1), 619–647.

——— (2018): "Railroads of the Raj: Estimating the impact of transportation infrastructure," *American Economic Review*, 108(4-5), 899–934.

Donaldson, D., and R. Hornbeck (2016): "Railroads and American economic growth: A "market access" approach," *The Quarterly Journal of Economics*, 131(2), 799–858.

Ducruet, C., R. Juhász, D. K. Nagy, C. Steinwender, et al. (2020): "All aboard: The effects of port development," Discussion paper.

Duranton, G., and M. A. Turner (2011): "The fundamental law of road congestion: Evidence from US cities," *American Economic Review*, 101(6), 2616–52.

Eaton, J., and S. Kortum (2002): "Technology, Geography and Trade," *Econometrica*, 70(5), 1741–1779.

——— (2012): "Putting Ricardo to work," *Journal of Economic Perspectives*, 26(2), 65–90.

Edwards, A. (2017): "U.S. Cities List," Discussion paper.

Eluru, N., A. Pinjari, J. Guo, I. Sener, S. Srinivasan, R. Copperman, and C. Bhat (2008): "Population updating system structures and models embedded in the comprehensive econometric microsimulator for urban systems," *Transportation Research Record: Journal of the Transportation Research Board*, (2076), 171–182.

EMC Research (2016): "2016 Center City Commuter Mode Split Survey," Discussion paper, Commute Seattle.

Fajgelbaum, P. D., and E. Schaal (2020): "Optimal transport networks in spatial equilibrium," *Econometrica*, 88(4), 1411–1452.

Fan, J., Y. Lu, and W. Luo (2019): "Valuing Domestic Transport Infrastructure: A View from the Route Choice of Exporters," .

Fan, J., and W. Luo (2020): "A Tractable Model of Transshipment," .

Federal Highway Administration (2015): "2015 Status of the Nation's Highways, Bridges, and Transit: Conditions and Performance," Discussion paper, US Department of Transportation.

——— (2016): "Highway Performance Monitoring System Field Manual," Discussion paper, Department of Transportation.

Feigenbaum, B., M. G. Fields, and S. Purnell (2020): "25th Annual Highway Report," *Reason Foundation*.

Feyrer, J. (2019): "Trade and income: Exploiting time series in geography," *American Economic Journal: Applied Economics*, 11(4), 1–35.

FHA (2008): "2008 status of the nation's highways, bridges, and transit: Conditions and perfomance," *Federal Highway Administration and Federal Transit Administration*.

Fogel, R. (1962): "A Quantitative Approach to the Study of Railroads in American Economic Growth: a Report of Some Preliminary Findings," *Journal of Economic History*, 22(2), 163–197.

Fogel, R. W. (1964): *Railroads and American economic growth: essays in econometric history*, vol. 296. Johns Hopkins Press Baltimore.

Galichon, A. (2016): *Optimal transport methods in economics*. Princeton University Press.

Ganapati, S., W. F. Wong, and O. Ziv (2020): "Entrepôt: Hubs, Scale, and Trade Costs," Discussion paper.

Holden, D. J. (1989): "Wardrop's third principle: urban traffic congestion and traffic policy," *Journal of Transport Economics and Policy*, pp. 239–262.

Hummels, D. L., and G. Schaur (2013): "Time as a trade barrier," *American Economic Review*, 103(7), 2935–59.

Kaszynski, W. (2000): *The American Highway: The History and Culture of Roads in the United States*. McFarland.

Lefevre, B., D. Leipziger, and M. Raifman (2014): "The Trillion Dollar Question: Tracking public and private investment in transport," *Washington, D. C: World Resources Institute*.

Mattsson, L.-G., J. W. Weibull, and P. O. Lindberg (2014): "Extreme values, invariance and choice probabilities," *Transportation Research Part B: Methodological*, 59, 81–95.

NSTIFC (2009): "Paying our way: A new framework for transportation finance," *National Surface Transportation Infrastructure Financing Commission*.

Office of the State Auditor, D. o. T. A. (2002): *Capitalization and Depreciation of Infrastructure*. Mississippi Association of Governmental Purchasing and Property Agents.

Osher, S., and J. Sethian (1988): "Fronts Propagating with Curvature-dependent Speed: Algorithms based on Hamilton-Jacobi Formulations," *Journal of computational physics*, 79(1), 12–49.

Pascali, L. (2017): "The wind of change: Maritime technology, trade, and economic development," *American Economic Review*, 107(9), 2821–54.

Polyanin, A., and A. Manzhirov (2008): *Handbook of Integral Equations*. Chapman & Hall/CRC.

Redding, S., and A. J. Venables (2004): "Economic geography and international inequality," *Journal of international Economics*, 62(1), 53–82.

Redding, S. J. (2016): "Goods trade, factor mobility and welfare," *Journal of International Economics*, 101, 148–167.

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

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

Rosenthal, S. S., and W. C. Strange (2004): "Evidence on the nature and sources of agglomeration economies," *Handbook of regional and urban economics*, 4, 2119–2171.

Sheffi, Y. (1985): *Urban transportation networks*, vol. 6. Prentice-Hall, Englewood Cliffs, NJ.

Szabo, F. (2015): *The linear algebra survival guide: illustrated with Mathematica*. Academic Press.

Tsitsiklis, J. (1995): "Efficient Algorithms for Globally Optimal Trajectories," *Automatic Control, IEEE Transactions on*, 40(9), 1528–1538.

Tsivanidis, N. (2018): "The aggregate and distributional effects of urban transit infrastructure: Evidence from bogotá's transmilenio," .

Vickrey, W. (1967): "Optimization of traffic and facilities," *Journal of Transport Economics and Policy*, pp. 123–136.

Washington State Legislature (1965): "RCW 46.61.400," Electronic.

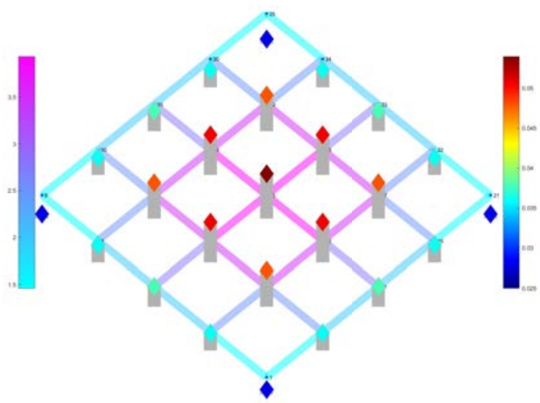
Weber, H. J., and G. B. Arfken (2003): *Essential Mathematical Methods for Physicists*. Academic Press.

Weber, S., and M. Péclat (2017): "GEOROUTE: Stata module to calculate travel distance and travel time between two addresses or two geographical points," .

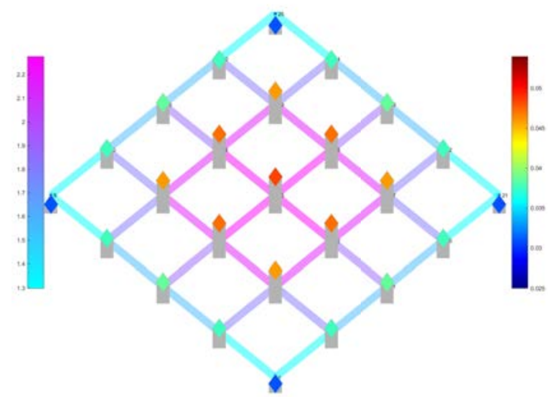
Tables and Figures

Figure 1: Traffic Congestion and the Distribution of Economic Activity

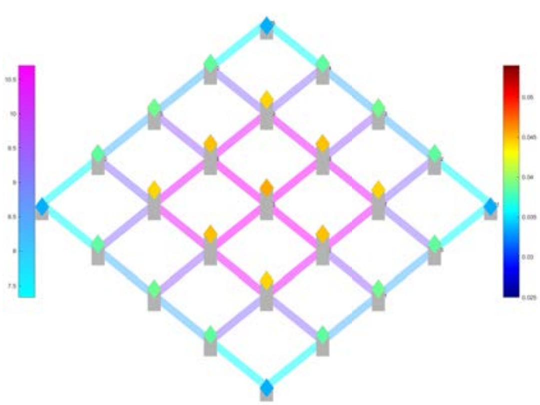
(a) No congestion ($\lambda = 0$, $\bar{L} = 100$)



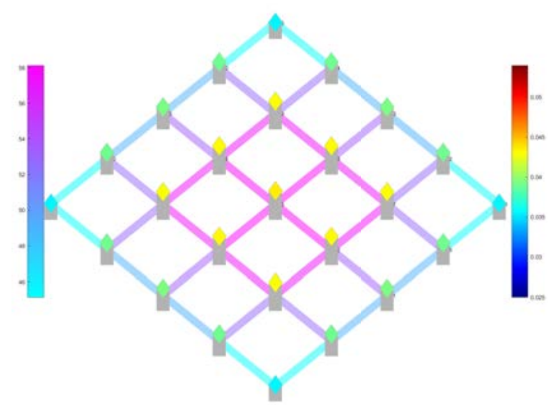
(b) Congestion, low scale ($\lambda = 0.05$, $\bar{L} = 100$)



(c) Congestion, medium scale ($\lambda = 0.05$, $\bar{L} = 1000$)

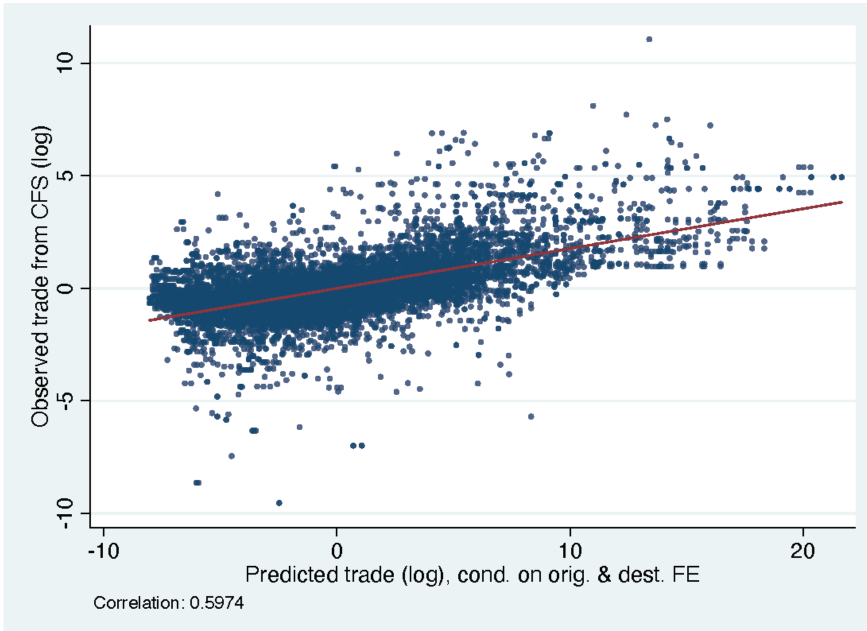


(d) Congestion, large scale ($\lambda = 0.05$, $\bar{L} = 10000$)

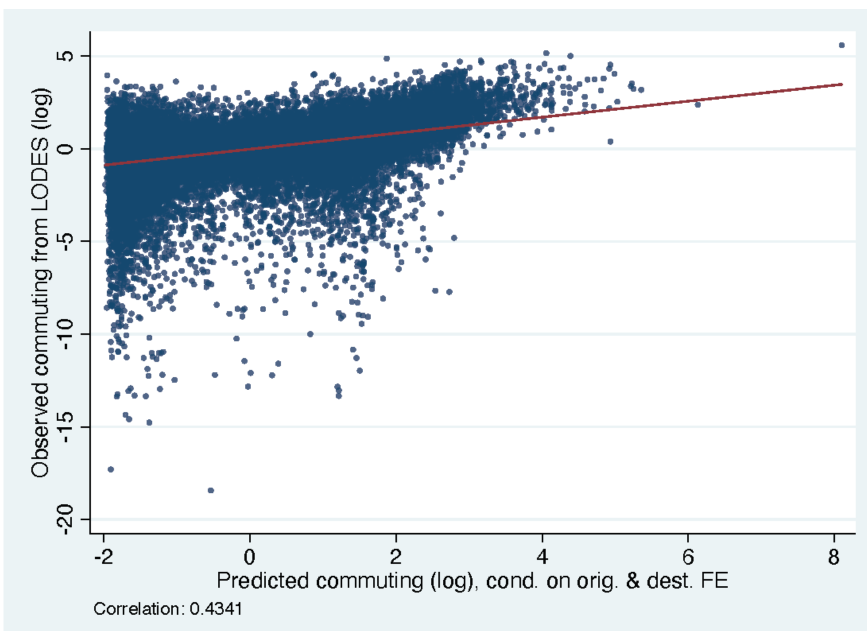


Notes: This figure shows how traffic congestion (λ) and the scale of the economy $\left(\bar{L}\right)$ shapes the distribution of economic activity within an example 5x5 grid network using the urban model. The height of the buildings (and the rooftop colors, associated with the color bar on the right) indicate the equilibrium residential population (L_i^R) at each location in the city, and the color of each link (associated with the color bar on the left) indicates the equilibrium traffic along the link. Throughout, $\alpha = \beta = 0$, $\theta = 4$, and $\bar{t}_{kl} = 1.5$ for connected links.

Figure 2: Predicting Flows Using Traffic



(a) Trade flows in an economic geography model



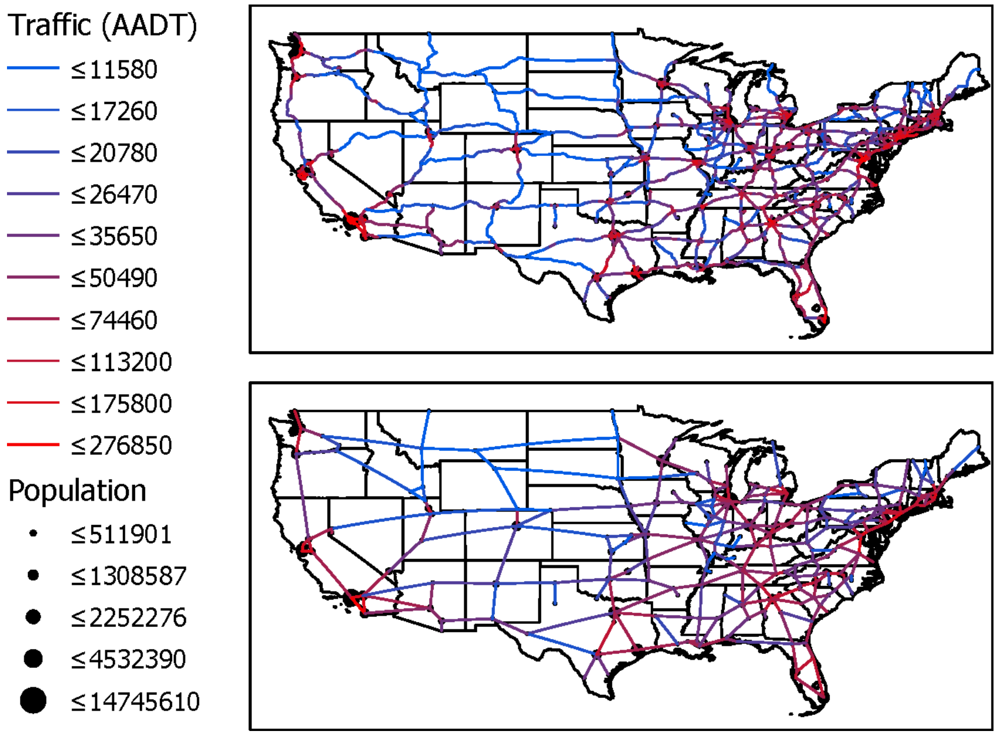
(b) Commuting flows in an urban model

Notes: This figure compares the observed bilateral origin to destination flows to those predicted from the observed traffic along the transportation network. In panel (a), we compare the predicted (log) trade flows on the x-axis to the observed (log) trade flows between metropolitan areas from the Commodity Flow Survey (CFS) data on the y-axis using the economic geography model. In panel (b), we compare the predicted (log) commuting flows on the x-axis to the observed (log) commuting flows from the Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES) between grid cells within Seattle. In both figures, the predicted and observed flows are residualized using origin and destination fixed effects, so the observed correlation only arises through similarity at the pair level.

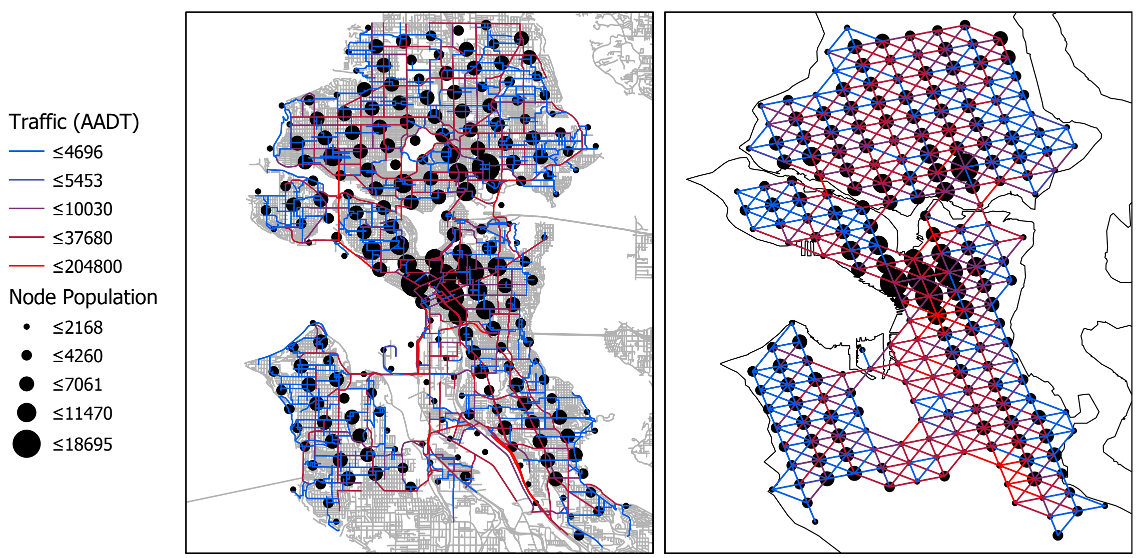
Figure 3: Transportation Systems and their Network Representations

Panel A: The Interstate Highway System

(a) U.S. Highway Network



(b) Seattle Road Network

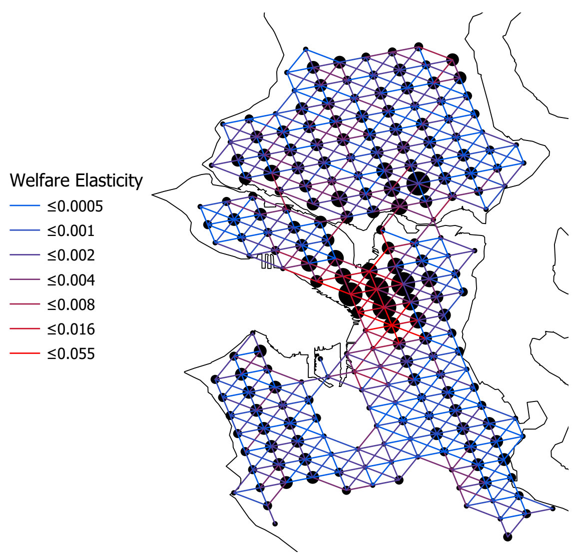


Notes: This figure presents the observed transportation network (on the top) and the constructed infrastructure matrix (on the bottom) for the U.S. highway network (panel a) and the observed transportation network (on the right) and the constructed infrastructure matrix (on the right) for the Seattle road network (panel b). In both panels, the size of each node reflects its population and the color of each link reflects the amount of traffic with red (blue) indicating high (low) levels of traffic. The gray roads in panel (b) are roads not on the least cost route between grid centers.

Figure 4: Welfare Elasticities of Infrastructure Improvement

(a) U.S. Highway Network

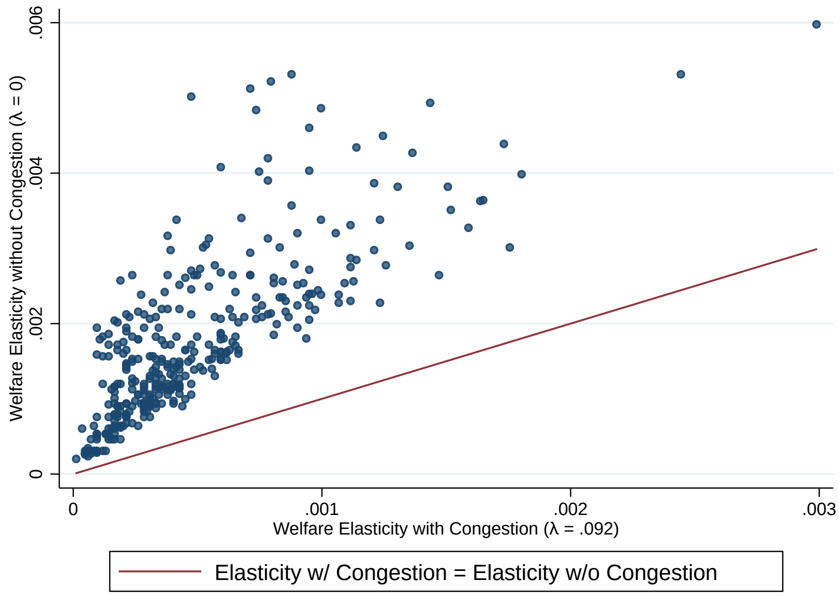
(b) Seattle Road Network



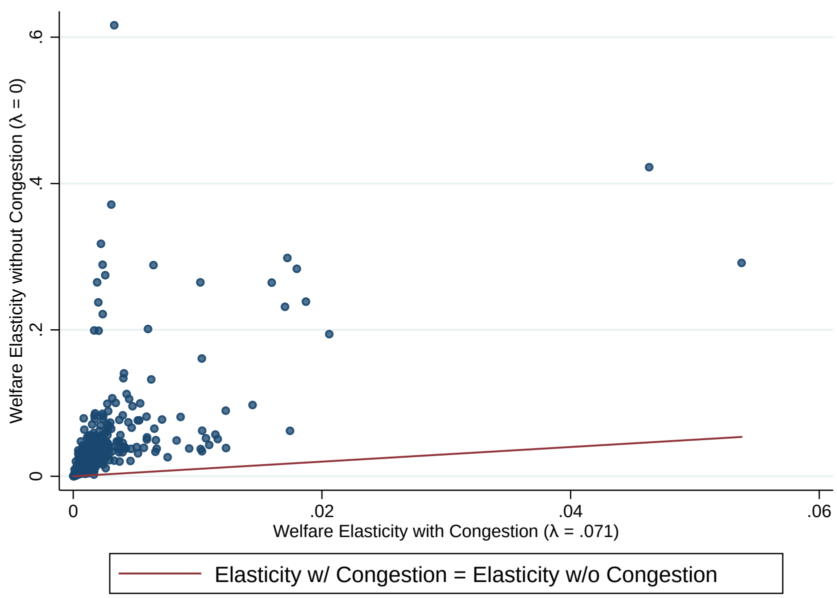
Notes: This figure presents the elasticity of aggregate welfare to improving each link in the U.S. Highway Network (panel A) and the Seattle road network (Panel B). The color ramp goes from blue (lower welfare elasticity) to red (higher welfare elasticity). Nodes in the network are marked by the black circles, which are increasing the population size of the node.

Figure 5: Comparing Welfare Elasticities With and Without Congestion

(a) U.S. Highway Network



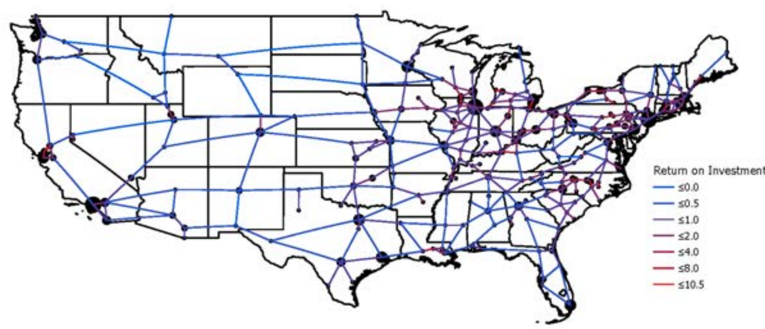
(b) Seattle Road Network



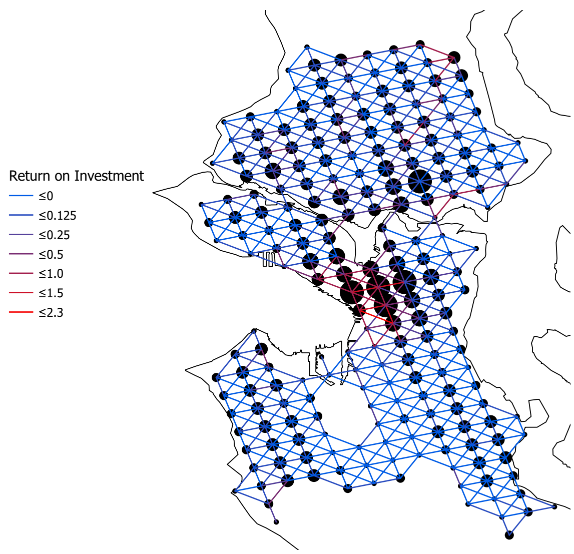
Notes: This figure compares the welfare elasticity calculated allowing for traffic congestion (given the estimated strength of congestion λ) to the welfare elasticity that would be calculated if traffic congestion were ignored (i.e. if $\lambda = 0$), as in a standard spatial model for each link in the U.S. highway network (panel a) and the Seattle road network (panel b).

Figure 6: Returns On Investment of Infrastructure Improvement

(a) U.S. Highway Network



(b) Seattle Road Network



Notes: This figure presents the return on investment of improving links in the Interstate Highway System (Panel A) and the Seattle road network (Panel B). Return on investment is annual and in decimals of the initial investment (i.e. 0.75 means a 75% return on initial investment per annum). The color ramps goes from blue (negative returns) to red (high positive returns). Nodes in the network are marked by the black circles, which are increasing the population size of the node.

Table 1: Estimating the strength of traffic congestion

Panel A: Interstate Highway System

	(1) OLS	(2) OLS	(3) IV: 1st stage	(4) IV: 2nd stage
Log congestion	0.109***	0.050***		0.739***
	(0.010)	(0.012)		(0.181)
Log distance			-0.156***	
			(0.033)	
Start-location FE	No	Yes	Yes	Yes
End-location FE	No	Yes	Yes	Yes
F-statistic	120.994	17.929	22.748	16.723
Observations (excl. singletons)	630	630	630	630
Observations (incl. singletons)	704	704	704	704

Panel B: Seattle Road Network

	(1) OLS	(2) IV: 1st stage	(3) IV	(4) IV: 1st stage	(5) IV
AADT per Lane	-0.048***		0.118**		0.488*
	(0.007)		(0.048)		(0.278)
Turns along Route		-0.252***		-0.091**	
		(0.049)		(0.039)	
Start-location FE	Yes	Yes	Yes	Yes	Yes
End-location FE	Yes	Yes	Yes	Yes	Yes
No. of Intersections	No	Yes	Yes	Yes	Yes
Bilateral Route Quality	No	No	No	Yes	Yes
F-statistic	41.546	26.347	6.191	5.336	3.084
Observations	1338	1338	1338	1338	1338

Notes: Panel A presents the congestion parameter estimates for the Interstate Highway System, and each observation is a segment of the interstate highway network. In columns 1, 2, and 4, the dependent variable is the (log) time of travel per unit distance, calculated using the HERE API, and the independent variable is the (log) AADT per lane from the highway performance monitoring system (HPMS). In column 3, we instrument for the (log) traffic per lane using the (log) length of the segment. Panel B presents the congestion parameter estimates for Seattle, and each observation is a segment of the Seattle's Network. In columns 1, 3, and 5 the dependent variable is the (log) time of travel per unit distance, calculated using the HERE API, and the independent variable is the (log) traffic per lane from the highway performance monitoring system (HPMS). In columns 3 and 5, we instrument for the (log) traffic per lane using the (log) number of turns, conditional on number of intersections traversed. In column 5, we add controls for bilateral route quality, which are generated by binning each segment into deciles based on the shares of arterial roads and local roads along it. Road classifications (functional system) are taken from the HPMS. For both panels, standard errors two-way clustered at the start-location and end-location are reported in parentheses. Stars indicate statistical significance: * $p<.10$ ** $p<.05$ *** $p<.01$.

Table 2: Return on investment rankings

Panel A: Interstate Highway System

	Origin	Destination	Rt. Number	RoI	Benefit ($m)	Cost ($m)
1	Kingsport-Bristol-Bristol, TN-VA	Johnson City, TN	I-26	10.43	55.492	5.271
2	Greensboro-High Point, NC	Winston-Salem, NC	I-40	9.54	158.548	16.455
3	Rochester, NY	Batavia, NY	I-90	7.31	117.779	15.896
4	Springfield, OH	Dayton, OH	I-70	6.38	175.536	27.07
5	Durham-Chapel Hill, NC	Raleigh-Cary, NC	I-40	5.95	274.062	45.313
6	Grand Rapids-Wyoming, MI	Muskegon-Norton Shores, MI	I-96	5.32	108.719	20.075
7	Toledo, OH	Monroe, MI	I-75	4.97	159.681	31.504
8	Vallejo-Fairfield, CA	Sacramento–Arden-Arcade–Roseville, CA	I-80	4.51	236.69	51.33
9	Breezewood, PA	Bedford, PA	I-76	4.27	41.902	9.598
10	Rockford, IL	Rochelle, IL	I-39	4.02	75.877	18.411

Panel B: Seattle Road Network

	Home Neighborhood	Home Address	Work Neighborhood	Work Address	RoI	Benefit ($m)	Cost ($m)
1	Downtown Seattle	805 S Jackson St	Downtown Seattle	83 Spring St	2.3	24.587	10.26
2	Downtown Seattle	83 Spring St	Downtown Seattle	1001 Columbia St	1.3	5.316	3.794
3	Downtown Seattle	805 S Jackson St	Downtown Seattle	1001 Columbia St	1.29	21.189	15.201
4	Capitol Hill	1698 Yale Ave	Capitol Hill	107A 13th Ave E	1.03	2.598	2.296
5	Downtown Seattle	2020 4th Ave	Capitol Hill	1698 Yale Ave	1.01	4.684	4.207
6	Downtown Seattle	2020 4th Ave	South Lake Union	401 Aurora Ave N	.91	3.474	3.44
7	Downtown Seattle	2020 4th Ave	Lower Queen Anne	201 Queen Anne Ave N	.89	7.973	8.024
8	Capitol Hill	1698 Yale Ave	South Lake Union	1214 E Roy St	.87	6.596	6.823
9	SoDo	562 1st Ave S	Downtown Seattle	83 Spring St	.79	3.021	3.404
10	Downtown Seattle	83 Spring St	Downtown Seattle	2020 4th Ave	.76	5.619	6.518
11	South Lake Union	401 Aurora Ave N	South Lake Union	1214 E Roy St	.75	4.738	5.545
12	North Seattle	2517 NE 120th St	Cedar Park	13504 WA-522	.75	2.026	2.381
13	Northeast Seattle	4200 Mary Gates Memorial Dr NE N207	Northeast Seattle	4700 45th Ave NE	.74	1.261	1.5
14	Downtown Seattle	1001 Columbia St	Downtown Seattle	2020 4th Ave	.73	5.001	6.052
15	Downtown Seattle	83 Spring St	Capitol Hill	1698 Yale Ave	.71	4.261	5.259
16	Greater Duwamish	1150 S Atlantic St	Downtown Seattle	805 S Jackson St	.67	9.416	12.306
17	SoDo	1705 4th Ave S	SoDo	562 1st Ave S	.66	2.696	3.567
18	Downtown Seattle	1001 Columbia St	Capitol Hill	1698 Yale Ave	.65	3.802	5.066
19	Cedar Park	13504 WA-522	Olympic Hills	14322 19th Ave NE	.65	2.373	3.177
20	Downtown Seattle	2020 4th Ave	South Lake Union	1214 E Roy St	.64	4.735	6.4

Notes: Names of the origin and destinations in Panel A are sourced from NH-GIS data on the location and area of CBSAs and from Google Maps. Names and addresses of the home and work locations in Panel B are sourced from the Google Maps Geocoding API. RoIs are reported as annualized decimal returns (i.e 1.2 means an annual 120% return). Benefits in Panel A are reported in millions of chains 2012 USD, and benefits in Panel B are reported in millions of 2016 USD.

A Appendix: Main Derivations

This section presents the full derivations of the major results mentioned in the text.

A.1 Section 3.1: Transportation Costs

Define the $N \times N$ matrix $\mathbf{A} = \left[a_{ij} \equiv t_{ij}^{-\theta}\right]$. We can write τ_{ij} from equation 3 by explicitly summing across all possible routes of all possible lengths. To do so, we sum across all locations that are traveled through all the possible paths as follows:

$$\tau_{ij}^{-\theta} = \sum_{K=0}^{\infty} \left(\sum_{k_1=1}^{N} \sum_{k_2=1}^{N} ... \sum_{k_{K-1}=1}^{N} a_{i,k_1} \times a_{k_1,k_2} \times ... \times a_{k_{K-2},k_{K-1}} \times a_{k_{K-1},j} \right),$$

where k_n is the sub-index for the n^{th} location arrived at on a particular route. Note that pairs of locations that are not connected will have $a_{ij} = 0$, so that infeasible routes do not affect the sum. The portion of the expression in the parentheses is equivalent to the (i, j) element of the weighted adjacency matrix to the power K, i.e.:

$$\tau_{ij}^{-\theta} = \sum_{K=0}^{\infty} \mathbf{A}_{ij}^{K},$$

where $\mathbf{A}^K = \left[A_{ij}^K\right]$, i.e. A_{ij}^K is the (i, j) element of the matrix $\mathbf{A}$ to the matrix power K. As we note in the paper, for a matrix $\mathbf{A}$ with spectral radius less than one, the geometric sum can be expressed as:

$$\sum_{K=0}^{\infty} \mathbf{A}^K = (\mathbf{I} - \mathbf{A})^{-1} \equiv \mathbf{B},$$

where $\mathbf{B} = [b_{ij}]$ is the Leontief inverse of the weighted adjacency matrix, so the transportation cost from i to j can be expressed as a function of the infrastructure matrix:

$$\tau_{ij}^{-\theta} = b_{ij},$$

as in equation (21).

A.2 Section 3.2: Traffic Flows

Beginning with equation (22) we have:

$$\begin{aligned}
\pi_{ij}^{kl} &= \sum_{r \in \Re_{ij}} \frac{\pi_{ij,r}}{\sum_{r' \in \Re_{ij}} \pi_{ij,r'}} n_r^{kl} \iff \\
\pi_{ij}^{kl} &= \sum_{r \in \Re_{ij}} \frac{\left(\prod_{l=1}^{K} t_{r_{l-1},r_l}^{-\theta}\right)}{\sum_{r \in \Re_{ij}} \left(\prod_{l=1}^{K} t_{r_{l-1},r_l}^{-\theta}\right)} n_r^{kl} \iff \\
\pi_{ij}^{kl} &= \tau_{ij}^{\theta} \sum_{r \in \Re_{ij}} \prod_{l=1}^{K} t_{r_{l-1},r_l}^{-\theta} n_r^{kl},
\end{aligned}$$

where the second line used either equation (1) (for the economic geography model) or equation (12) (for the urban model), and the third line used the definition of τ_{ij} from equation (3).

For each route in $r \in \Re_{ij}$, the value $\prod_{l=1}^{K} t_{r_{l-1},r_l}^{-\theta} n_r^{kl}$ is the transportation costs incurred along the route multiplied by the number of times the routes traverses link $\{k, l\}$. To calculate this, we proceed by summing across all possible traverses that occur on all routes from i to j. To do so, note for any $r \in \Re_{ij}$ of length K (the set of which we denote as $\Re_{ij,K}$), a traverse is possible at any point $B \in [1, 2, ..., K-1]$ in the route.

Defining $\mathbf{A} \equiv [a_{kl}] = \left[t_{kl}^{-\theta}\right]$ and $\mathbf{B} \equiv [b_{ij}] = \left[\tau_{ij}^{-\theta}\right]$ as above, we can write:

$$\pi_{ij}^{kl} = \frac{1}{b_{ij}} \sum_{K=0}^{\infty} \sum_{B=0}^{K-1} \left(\sum_{r \in \Re_{ik,B}} \prod_{n=1}^{B} a_{r_{n-1},r_n} \right) \times a_{kl} \times \left(\sum_{r \in \Re_{kj,K-B-1}} \prod_{n=1}^{K-B-1} a_{r_{n-1},r_n} \right)$$

This can in turn allows us to explicitly enumerate all possible paths from i to k of length B and all possible paths from l to j of length $K - B - 1$:

$$\pi_{ij}^{kl} = \frac{1}{b_{ij}} \sum_{K=0}^{\infty} \sum_{B=0}^{K-1} \left(\sum_{n_1=1}^{N} \cdots \sum_{n_{B-1}=1}^{N} a_{i,n_1} \times \ldots \times a_{n_{B-1},k} \right) \times a_{kl} \times \left(\sum_{n_1=1}^{N} \cdots \sum_{n_{K-B-1}=1}^{N} a_{l,n_1} \times \ldots \times a_{n_{K-B-1},j} \right),$$

which can be expressed more succinctly as elements of matrix powers of $\mathbf{A}$:

$$\pi_{ij}^{kl} = \frac{1}{b_{ij}} \sum_{K=0}^{\infty} \sum_{B=0}^{K-1} A_{ik}^{B} \times a_{kl} \times A_{lj}^{K-B-1}.$$

A result from matrix calculus (see e.g. Weber and Arfken (2003)) is for any $N \times N$ matrix $\mathbf{C}$ we have:

$$\sum_{K=0}^{\infty} \sum_{B=0}^{K-1} \mathbf{A}^{B} \mathbf{C} \mathbf{A}^{K-B-1} = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{C} (\mathbf{I} - \mathbf{A})^{-1}. \tag{A.1}$$

Define $\mathbf{C}$ to be an $N \times N$ matrix that takes the value of a_{kl} at row k and column l and zeros everywhere else. Using equation (A.1) we obtain our result:

$$\begin{aligned} \pi_{ij}^{kl} &= \frac{b_{ik} a_{kl} b_{lj}}{\tau_{ij}^{-\theta}} \iff \\ \pi_{ij}^{kl} &= \frac{\tau_{ik}^{-\theta} t_{kl}^{-\theta} \tau_{lj}^{-\theta}}{\tau_{ij}^{-\theta}}, \end{aligned} \tag{A.2}$$

as in equation (23).

We now derive gravity equations for traffic over a link for both economic geography and commuting models. For trade, we sum over all trade between all origins and destinations, and all routes taken by that trade, to get:

$$\begin{aligned} \Xi_{kl} &= \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \sum_{r \in \Re_{ij}} \pi_{ij,r} n_r^{kl} E_j \iff \\ \Xi_{kl} &= \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \pi_{ij}^{kl} X_{ij} \iff \\ \Xi_{kl} &= \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \frac{\tau_{ik}^{-\theta} t_{kl}^{-\theta} \tau_{lj}^{-\theta}}{\tau_{ij}^{-\theta}} \times \tau_{ij}^{-\theta} \frac{Y_i}{\Pi_i^{-\theta}} \frac{E_j}{P_j^{-\theta}} \iff \\ \Xi_{kl} &= t_{kl}^{-\theta} \sum_{i \in \mathcal{N}} \tau_{ik}^{-\theta} \frac{Y_i}{\Pi_i^{-\theta}} \sum_{j \in \mathcal{N}} \tau_{lj}^{-\theta} \frac{E_j}{P_j^{-\theta}}, \end{aligned}$$

where the second line used equations (1) and (22), the third lines used equation (23), and the fourth lined rearranged. Recalling our definition of the consumer and producer market access terms (equations (7) and (6)) in the text, this becomes:

$$\Xi_{kl} = t_{kl}^{-\theta} \times P_k^{-\theta} \times \Pi_l^{-\theta},$$

as in equation (24).

Turning to the commuting model, we proceed similarly, summing over all commuting flow pairs and the routes they take:

$$
\begin{aligned}
\Xi_{kl} &= \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}}\sum_{r\in\Re_{ij}} \pi_{ij,r} n_r^{kl} \bar{L} \iff \\
\Xi_{kl} &= \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} \pi_{ij}^{kl} L_{ij} \iff \\
\Xi_{kl} &= \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} \frac{\tau_{ik}^{-\theta} t_{kl}^{-\theta} \tau_{lj}^{-\theta}}{\tau_{ij}^{-\theta}} \times \tau_{ij}^{-\theta} \frac{L_i^R}{\Pi_i^{-\theta}} \frac{L_j^F}{P_j^{-\theta}} \iff \\
\Xi_{kl} &= \frac{\bar{L}}{W^\theta} \times t_{kl}^{-\theta} \times \left(\sum_{i\in\mathcal{N}} \tau_{ik}^{-\theta} \frac{L_i^R}{\Pi_i^{-\theta}}\right) \times \left(\sum_{j\in\mathcal{N}} \tau_{lj}^{-\theta} \frac{L_j^F}{P_j^{-\theta}}\right),
\end{aligned}
$$

where the second line used equations (12) and and (22), the third lines used equation (23), and the fourth lined rearranged. We substitute in the consumer and producer market access defined in (16) and (17) to generate traffic gravity for the commuting framework:

$$
\Xi_{kl} = t_{kl}^{-\theta} \times (P_k)^{-\theta} \times (\Pi_l)^{-\theta},
$$

again as in equation (24).

A.3 Section 5.1: Traffic, Trade, and Commuting flows

In this section, we derive an analytical mapping between traffic and gravity flows of trade and commuting. We begin by writing trade, commuting and traffic gravity from equations (2), (13), and (24), respectively in matrix form:

$$
\begin{aligned}
\mathbf{X} &= \left(\frac{Y}{\Pi}\right)(\mathbf{I}-\mathbf{A})^{-1}\left(\frac{E}{P}\right) \\
\mathbf{L} &= \left(\frac{L^R}{\Pi}\right)(\mathbf{I}-\mathbf{A})^{-1}\left(\frac{L^F}{P}\right) \\
\boldsymbol{\Xi} &= \mathbf{P}\mathbf{A}\boldsymbol{\Pi},
\end{aligned}
$$

where for the gravity equations we used equation (21) and where

$$
\begin{aligned}
\mathbf{X} &= [X_{ij}] & \mathbf{L} &= [L_{ij}] \\
\boldsymbol{\Xi} &= [\Xi_{ij}] & \mathbf{A} &= \left[t_{ij}^{-\theta}\right] \\
\mathbf{P} &= \text{diag}\left(P_i^{-\theta}\right) & \boldsymbol{\Pi} &= \text{diag}\left(\Pi_i^{-\theta}\right) \\
Y &= \text{diag}\left(Y_i\right) & E &= \text{diag}(E_i) \\
L^R &= \text{diag}\left(L_i^R\right) & L^F &= \text{diag}\left(L_i^F\right) \\
\frac{Y}{\Pi} &= \text{diag}\left(\frac{Y_i}{\Pi_i^{-\theta}}\right) & \frac{E}{P} &= \text{diag}\left(\frac{E_i}{P_i^{-\theta}}\right) \\
\frac{L^R}{\Pi} &= \text{diag}\left(\frac{L_i^R}{\Pi_i^{-\theta}}\right) & \frac{L^F}{P} &= \text{diag}\left(\frac{L_i^F}{P_i^{-\theta}}\right)
\end{aligned}
$$

are each $N \times N$ matrices.

Solving for the adjacency matrix $\mathbf{A}$, we have:

$$\begin{aligned}\boldsymbol{\Xi} &= \mathbf{P}\mathbf{A}\boldsymbol{\Pi} \iff \\ \mathbf{A} &= \mathbf{P^{-1}}\boldsymbol{\Xi}\boldsymbol{\Pi^{-1}}\end{aligned}$$

which we can substitute into our trade gravity equation:

$$\begin{aligned}\mathbf{X} &= \left(\frac{Y}{\boldsymbol{\Pi}}\right)(\mathbf{I}-\mathbf{A})^{-1}\left(\frac{E}{\mathbf{P}}\right) \iff \\ \mathbf{X} &= \left(\frac{Y}{\Pi}\right)\left(\mathbf{I}-\mathbf{P^{-1}}\boldsymbol{\Xi}\boldsymbol{\Pi^{-1}}\right)^{-1}\left(\frac{E}{\mathbf{P}}\right) \iff \\ \mathbf{X}^{-1} &= \left(\frac{E}{\mathbf{P}}\right)^{-1}\left(\mathbf{I}-\mathbf{P^{-1}}\boldsymbol{\Xi}\boldsymbol{\Pi^{-1}}\right)\left(\frac{Y}{\Pi}\right)^{-1} \iff \\ \mathbf{X}^{-1} &= \left(\frac{E}{\mathbf{P}}\right)^{-1}\left(\frac{Y}{\boldsymbol{\Pi}}\right)^{-1} - (E)^{-1}\,\Xi\,(Y)^{-1} \iff \\ \mathbf{X}^{-1} &= (E)^{-1}\,(\mathbf{P}\boldsymbol{\Pi}-\Xi)\,(Y)^{-1} \iff \\ \mathbf{X} &= (Y)\,(\mathbf{P}\boldsymbol{\Pi}-\Xi)^{-1}\,(E)\end{aligned}$$

and our commuting gravity equation:

$$\begin{aligned}\mathbf{L} &= \left(\frac{L^R}{\Pi}\right)(\mathbf{I}-\mathbf{A})^{-1}\left(\frac{L^F}{P}\right) \iff \\ \mathbf{L} &= \left(\frac{L^R}{\Pi}\right)\left(\mathbf{I}-\mathbf{P^{-1}}\boldsymbol{\Xi}\boldsymbol{\Pi^{-1}}\right)^{-1}\left(\frac{L^F}{P}\right) \iff \\ \mathbf{L}^{-1} &= \left(\frac{L^F}{P}\right)^{-1}\left(\mathbf{I}-\mathbf{P^{-1}}\boldsymbol{\Xi}\boldsymbol{\Pi^{-1}}\right)\left(\frac{L^R}{\Pi}\right)^{-1} \iff \\ \mathbf{L}^{-1} &= \left(\left(\frac{L^F}{P}\right)^{-1}\left(\frac{L^R}{\Pi}\right)^{-1} - \left(L^f\right)^{-1}\boldsymbol{\Xi}\left(L^R\right)^{-1}\right) \iff \\ \mathbf{L}^{-1} &= \left(L^F\right)^{-1}(\mathbf{P}\boldsymbol{\Pi}-\boldsymbol{\Xi})\left(L^R\right)^{-1} \iff \\ \mathbf{L} &= \left(L^R\right)(\mathbf{P}\boldsymbol{\Pi}-\boldsymbol{\Xi})^{-1}\left(L^F\right)\end{aligned}$$

Now all that remains is to define diagonal matrix $\mathbf{P}\boldsymbol{\Pi}$ in terms of traffic Ξ and other observables. For the trade model, we have the following, where P and $\frac{Y}{\Pi}$ are column vectors:

$$\begin{aligned}P_i^{-\theta} &= \sum_j \tau_{ji}^{-\theta}\frac{Y_j}{\Pi_j^{-\theta}} \iff \\ P &= \left(I-A^T\right)^{-1}\left(\frac{Y}{\Pi}\right) \iff \\ \left(I-\Pi^{-1}\Xi^T P^{-1}\right)P &= \frac{Y}{\Pi} \iff \\ P-\Pi^{-1}\Xi^T 1 &= \frac{Y}{\Pi} \\ \Pi P &= Y+\Xi^T 1\end{aligned}$$

and:

$$
\begin{aligned}
\Pi_i^{-\theta} &= \sum_j \tau_{ij}^{-\theta} \frac{E_j}{P_j^{-\theta}} \iff \\
\Pi &= (I - A)^{-1} \left(\frac{E}{P}\right) \iff \\
\left(I - P^{-1}\Xi\Pi^{-1}\right)\Pi &= \left(\frac{E}{P}\right) \\
\Pi - P^{-1}\Xi 1 &= \frac{E}{P} \\
\Pi P &= E + \Xi 1
\end{aligned}
$$

where Π and $\frac{E}{P}$ are column vectors.

Since we have two definitions of column vector ΠP, we average them:

$$
\Pi P = \frac{1}{2}(E + \Xi 1) + \frac{1}{2}\left(Y + \Xi^T 1\right) = \frac{1}{2}(E + Y) + \frac{1}{2}\left(\Xi 1 + \Xi^T 1\right)
$$

and plug that definition into our matrix product characterization of trade flows, where the diagonal matrix $\mathbf{P\Pi} = \mathrm{diag}(\Pi P)$:

$$
\begin{aligned}
\mathbf{X} &= (Y)\left(\mathrm{diag}\left(\frac{1}{2}(E + Y) + \frac{1}{2}\left(\Xi 1 + \Xi^T 1\right)\right) - \Xi\right)^{-1}(E) \iff \\
X_{ij} &= \left[\mathbf{D}^X - \Xi\right]_{ij}^{-1} \times Y_i \times E_j,
\end{aligned}
$$

as in equation (34).

For the commuting model, we derive the following for vector $P\Pi$:

$$
\begin{aligned}
P_i^{-\theta} &= \sum_j \tau_{ji}^{-\theta} \frac{L_j^R}{\Pi_j^{-\theta}} \iff \\
P &= \left(I - A^T\right)^{-1} \left(\frac{L^R}{\Pi}\right) \iff \\
\left(I - \Pi^{-1}\Xi^T P^{-1}\right) P &= \frac{L^R}{\Pi} \iff \\
P - \Pi^{-1}\Xi^T 1 &= \frac{L^R}{\Pi} \iff \\
\Pi P &= \frac{L^R}{\Pi} + \Xi^T 1
\end{aligned}
$$

where P and $\frac{L^R}{\Pi}$ are column vectors. We also derive another definition:

$$
\begin{aligned}
\Pi_i^{-\theta} &= \sum_j \tau_{ij}^{-\theta} \frac{L_j^F}{P_j^{-\theta}} \iff \\
\Pi &= (I - A)^{-1} \left(\frac{L^F}{P} \right) \iff \\
\left(I - P^{-1} \Xi \Pi^{-1} \right) \Pi &= \left(\frac{L^F}{P} \right) \\
\Pi - P^{-1} \Xi 1 &= \frac{L^F}{P} \\
\Pi P &= L^F + \Xi 1
\end{aligned}
$$

where Π and $\frac{L^F}{P}$ are column vectors.

Like with the trade case, we average over the two different definitions for traffic to define the vector ΠP:

$$
\Pi P = \frac{1}{2} \left(L^F + \Xi 1 \right) + \frac{1}{2} \left(L^R + \Xi^T 1 \right) = \frac{1}{2} \left(L^F + L^R \right) + \frac{1}{2} \left(\Xi 1 + \Xi^T 1 \right)
$$

and combining this average definition with our matrix product characterization of commuting flows, with the equality $\mathbf{P\Pi} = \text{diag}(\Pi P)$we get:

$$
\begin{aligned}
\mathbf{L} &= \left(L^R \right) \left(\text{diag} \left(\frac{1}{2} \left(L^F + L^R \right) + \frac{1}{2} \left(\Xi 1 + \Xi^T 1 \right) \right) - \mathbf{\Xi} \right)^{-1} \left(L^F \right) \\
L_{ij} &= \left[\mathbf{D}^L - \Xi \right]_{ij}^{-1} \times L_i^R \times L_j^F,
\end{aligned}
$$

as in equation (35).

B Appendix: Proof of Proposition 1

B.1 Preliminaries

In this subsection, we show how the economic geography and urban models both are special cases of the following mathematical system:

$$
x_{i,1} = D_{i,1} \frac{C_{i,1} x_{i,1}^{b_{11}} x_{i,2}^{b_{12}}}{\sum_{j \in \mathcal{N}} C_{j,1} x_{j,1}^{b_{11}} x_{j,2}^{b_{12}}} x_{i,1}^{\frac{a}{a+1}} x_{i,2}^{-\frac{1}{a+1}} + \sum_{j \in \mathcal{N}} K_{ij} x_{j,1}^{\frac{1}{a+1}} x_{j,2}^{-\frac{a}{a+1}} \;\; \forall i \in \mathcal{N} \tag{B.1}
$$

$$
x_{i,2} = D_{i,2} \frac{C_{i,2} x_{i,1}^{b_{21}} x_{i,2}^{b_{22}}}{\sum_{j \in \mathcal{N}} C_{j,2} x_{j,1}^{b_{21}} x_{j,2}^{b_{22}}} x_{i,1}^{-\frac{1}{a+1}} x_{i,2}^{\frac{a}{a+1}} + \sum_{j \in \mathcal{N}} K_{ji} x_{j,1}^{-\frac{a}{a+1}} x_{j,2}^{\frac{1}{a+1}} \;\; \forall i \in \mathcal{N} \tag{B.2}
$$

In the following sections, we then prove the existence and provide conditions for the uniqueness of any equilibrium characterized by equations (B.1) and (B.2).

To begin, note that gravity equation in both frameworks can be written as follows:

$$
F_{ij} = \tau_{ij}^{-\theta} \times \frac{\gamma_i}{\Pi_i^{-\theta}} \times \frac{\delta_j}{P_j^{-\theta}}, \tag{B.3}
$$

where in the economic geography model $F_{ij} \equiv X_{ij}$, $\gamma_i \equiv Y_i$, and $\delta_j \equiv E_j$ and in the urban model $F_{ij} \equiv L_{ij}$, $\gamma_i \equiv L_i^R$, and $\delta_j \equiv L_i^F$. Written like this, the market market clearing conditions in both frameworks can be expressed identically as follows:

$$\gamma_i = \sum_j F_{ij} \tag{B.4}$$

$$\delta_j = \sum_i F_{ij} \tag{B.5}$$

Substituting the gravity equation (B.3) into the two market clearing conditions yields:

$$\Pi_i^{-\theta} = \sum_j \tau_{ij}^{-\theta} \times \frac{\delta_j}{P_j^{-\theta}}$$

$$P_j^{-\theta} = \sum_i \tau_{ij}^{-\theta} \times \frac{\gamma_i}{\Pi_i^{-\theta}}$$

Substituting in the expression for the endogenous trade costs $\tau_{ij}^{-\theta}$ as a function of the transportation network from equation (21) and inverting each linear equation, the system becomes:

$$P_i^{-\theta}\Pi_i^{-\theta} = \delta_i + \sum_j t_{ij}^{-\theta} P_i^{-\theta} \Pi_j^{-\theta}$$

$$P_i^{-\theta}\Pi_i^{-\theta} = \gamma_i + \sum_j t_{ji}^{-\theta} P_j^{-\theta} \Pi_i^{-\theta}$$

Finally, expressing the transportation network as a function of equilibrium traffic and the infrastructure network from equation (26):

$$\left(P_i^{-\theta}\right)^{\frac{\theta\lambda}{1+\theta\lambda}} \Pi_i^{-\theta} = \delta_i \left(P_i^{-\theta}\right)^{-\frac{1}{1+\theta\lambda}} + \sum_j \left(\bar{t}_{ij}^{\frac{1-\theta}{1+\theta\lambda}}\right) \left(\Pi_j^{-\theta}\right)^{\frac{1}{1+\theta\lambda}}$$

$$P_i^{-\theta} \left(\Pi_i^{-\theta}\right)^{\frac{\theta\lambda}{1+\theta\lambda}} = \gamma_i \left(\Pi_i^{-\theta}\right)^{-\frac{1}{1+\theta\lambda}} + \sum_j \bar{t}_{ji}^{-\frac{\theta}{1+\theta\lambda}} \left(P_j^{-\theta}\right)^{\frac{1}{1+\theta\lambda}},$$

or, by defining $p_i \equiv P_i^{-\theta}$, $\pi_i \equiv \Pi_i^{-\theta}$, $a \equiv \frac{\theta\lambda}{1+\theta\lambda}$, and $K_{ij} \equiv \bar{t}_{ij}^{-\frac{\theta}{1+\theta\lambda}}$, we can write this more succinctly as:

$$p_i^a \pi_i = \delta_i p_i^{-(1-a)} + \sum_j K_{ij} \pi_j^{(1-a)} \tag{B.6}$$

$$p_i \pi_i^a = \gamma_i \pi_i^{-(1-a)} + \sum_j K_{ji} p_j^{(1-a)}. \tag{B.7}$$

We proceed by one final change of variables to get equations (B.6) and (B.7) into a form more amenable to define an operator to establish existence and uniqueness. Define $x_{i,1} \equiv p_i^a \pi_i$ and $x_{i,2} \equiv p_i \pi_i^a$. Note this in turn implies $p_i = x_{i,1}^{\frac{a}{a^2-1}} x_{i,2}^{-\frac{1}{a^2-1}}$ and $\pi_i = x_{i,1}^{-\frac{1}{a^2-1}} x_{i,2}^{\frac{a}{a^2-1}}$, so that equations (B.6) and (B.7) become:

$$x_{i,1} = \delta_i x_{i,1}^{\frac{a}{a+1}} x_{i,2}^{-\frac{1}{a+1}} + \sum_j K_{ij} x_{j,1}^{\frac{1}{a+1}} x_{j,2}^{-\frac{a}{a+1}} \tag{B.8}$$

$$x_{i,2} = \gamma_i x_{i,1}^{-\frac{1}{a+1}} x_{i,2}^{\frac{a}{a+1}} + \sum_j K_{ji} x_{j,1}^{-\frac{a}{a+1}} x_{j,2}^{\frac{1}{a+1}}. \tag{B.9}$$

The final step is to write $\{\gamma_i, \delta_i\}$ as functions of $\{x_{i,1}, x_{i,2}\}$. As mentioned in the text, this mapping between the endogenous economic activity and the market access variables differs depending on the model considered, so we do it separately for each.

B.1.1 The Economic Geography Model

In the economic geography model in equilibrium, we have $\gamma_i = \delta_i = Y_i$. From welfare equalization equation (8) we have:

$$P_i = \frac{1}{\bar{W}} \bar{u}_i L_i^{\beta-1} Y_i$$

and from the definition of the producer price index equation (6) we have:

$$\Pi_i = \bar{A}_i L_i^{1+\alpha} Y_i^{-\frac{\theta+1}{\theta}}.$$

Combining these two equations yields a log-linear system, which can be inverted to write Y_i as a function of p_i and π_i, as follows:

$$\begin{pmatrix} \ln p_i \\ \ln \pi_i \end{pmatrix} = \begin{pmatrix} \theta \ln \bar{W} - \theta \ln \bar{u}_i \\ -\theta \ln \bar{A} \end{pmatrix} + \begin{pmatrix} \theta(1-\beta) & -\theta \\ -\theta(1+\alpha) & (1+\theta) \end{pmatrix} \begin{pmatrix} \ln L_i \\ \ln Y_i \end{pmatrix} \iff$$
$$\begin{pmatrix} \ln L_i \\ \ln Y_i \end{pmatrix} = \begin{pmatrix} \theta(1-\beta) & -\theta \\ -\theta(1+\alpha) & (1+\theta) \end{pmatrix}^{-1} \begin{pmatrix} \ln p_i - \theta \ln \bar{W} + \theta \ln \bar{u}_i \\ \ln \pi_i + \theta \ln \bar{A} \end{pmatrix} \iff$$
$$\begin{pmatrix} \ln L_i \\ \ln Y_i \end{pmatrix} = \begin{pmatrix} \frac{\left(\frac{1+\theta}{\theta}\right)}{1-\beta-\theta(\alpha+\beta)} & \frac{1}{1-\beta-\theta(\alpha+\beta)} \\ \frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)} & \frac{(1-\beta)}{1-\beta-\theta(\alpha+\beta)} \end{pmatrix} \begin{pmatrix} \ln p_i - \theta \ln \bar{W} + \theta \ln \bar{u}_i \\ \ln \pi_i + \theta \ln \bar{A} \end{pmatrix},$$

so that:

$$Y_i = p_i^{\frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}} \pi_i^{\frac{(1-\beta)}{1-\beta-\theta(\alpha+\beta)}} \left(\frac{\bar{u}_i}{\bar{W}}\right)^{\theta \frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}} \bar{A}_i^{\theta \frac{(1-\beta)}{1-\beta-\theta(\alpha+\beta)}}.$$

We can then express Y_i as a function of $\{x_{1,i}, x_{2,i}\}$ as follows:

$$Y_i = x_{i,1}^{\frac{(a(1+\alpha)-(1-\beta))}{(a^2-1)(1-\beta-\theta(\alpha+\beta))}} x_{i,2}^{\frac{a(1-\beta)-(1+\alpha)}{(a^2-1)(1-\beta-\theta(\alpha+\beta))}} \left(\frac{\bar{u}_i}{\bar{W}}\right)^{\theta \frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}} \bar{A}_i^{\theta \frac{(1-\beta)}{1-\beta-\theta(\alpha+\beta)}},$$

or, more succinctly, as:

$$Y_i = W^{\rho} C_i x_{i,1}^{b_1} x_{i,2}^{b_2}, \tag{B.10}$$

where $b_1 \equiv \frac{a(1+\alpha)-(1-\beta)}{(a^2-1)(1-\beta-\theta(\alpha+\beta))}$, $b_2 \equiv \frac{a(1-\beta)-(1+\alpha)}{(a^2-1)(1-\beta-\theta(\alpha+\beta))}$, $C_i \equiv \bar{u}_i^{\theta \frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}} \bar{A}_i^{\theta \frac{(1-\beta)}{1-\beta-\theta(\alpha+\beta)}}$, and $\rho \equiv -\theta \frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}$. Substituting equation (B.10) into the equilibrium system (B.8) and (B.9) yields:

$$x_{i,1} = W^{\rho} C_i x_{i,1}^{\frac{a}{a+1}+b_1} x_{i,2}^{-\frac{1}{a+1}+b_2} + \sum_j K_{ij} x_{j,1}^{\frac{1}{a+1}} x_{j,2}^{-\frac{a}{a+1}} \tag{B.11}$$

$$x_{i,2} = W^{\rho} C_i x_{i,1}^{-\frac{1}{a+1}+b_1} x_{i,2}^{\frac{a}{a+1}+b_2} + \sum_j K_{ji} x_{j,1}^{-\frac{a}{a+1}} x_{j,2}^{\frac{1}{a+1}}. \tag{B.12}$$

Equations (B.11) and (B.12) will form the basis of the uniqueness analysis below (where we additionally impose a symmetric infrastructure matrix). For the analysis of existence, we proceed one step further and impose the normalization that $\sum_i Y_i = \bar{L}$,which then allows us to write:

$$Y_i = \bar{L} \frac{C_i x_{i,1}^{b_1} x_{i,2}^{b_2}}{\sum_j C_j x_{j,1}^{b_1} x_{j,2}^{b_2}}$$

so that the system of equations becomes:

$$x_{i,1} = \bar{L} \frac{C_i x_{i,1}^{b_1} x_{i,2}^{b_2}}{\sum_j C_j x_{j,1}^{b_1} x_{j,2}^{b_2}} x_{i,1}^{\frac{a}{a+1}} x_{i,2}^{-\frac{1}{a+1}} + \sum_j K_{ij} x_{j,1}^{\frac{1}{a+1}} x_{j,2}^{-\frac{a}{a+1}} \tag{B.13}$$

$$x_{i,2}=\bar{L}\frac{C_i x_{i,1}^{b_1}x_{i,2}^{b_2}}{\sum_j C_j x_{j,1}^{b_1}x_{j,2}^{b_2}}x_{i,1}^{-\frac{1}{a+1}}x_{i,2}^{\frac{a}{a+1}}+\sum_j K_{ji}x_{j,1}^{-\frac{a}{a+1}}x_{j,2}^{\frac{1}{a+1}}, \tag{B.14}$$

which is a special case of equations (B.1) and (B.2), as claimed.

B.1.2 The Urban Model

In the urban model, we have $\gamma_i \equiv L_i^R$ and $\delta_i \equiv L_i^F$ and, from equations (16) and (17) we have:

$$P_i^{-\theta}=\bar{A}_i^{-\theta}\left(\frac{\bar{L}}{\bar{W}^\theta}\right)^{\frac{1}{2}}\left(L_i^F\right)^{1-\theta\alpha}$$

$$\Pi_i^{-\theta}=\bar{u}_i^{-\theta}\left(\frac{\bar{L}}{\bar{W}^\theta}\right)^{\frac{1}{2}}\left(L_i^R\right)^{1-\theta\beta},$$

so that $L_i^F=p_i^{\frac{1}{1-\theta\alpha}}\bar{A}_i^{\frac{\theta}{1-\theta\alpha}}\left(\frac{\bar{W}^\theta}{\bar{L}}\right)^{\frac{1}{2(1-\theta\alpha)}}$ and $L_i^R=\pi_i^{\frac{1}{1-\theta\beta}}\bar{u}_i^{\frac{\theta}{1-\theta\beta}}\left(\frac{\bar{W}^\theta}{\bar{L}}\right)^{\frac{1}{2(1-\theta\beta)}}$. We can then express L_i^F and L_i^R as a function of $\{x_{1,i},x_{2,i}\}$ as follows:

$$L_i^F=x_{i,1}^{\frac{a}{a^2-1}\frac{1}{1-\theta\alpha}}x_{i,2}^{-\frac{1}{a^2-1}\frac{1}{1-\theta\alpha}}\bar{A}_i^{\frac{\theta}{1-\theta\alpha}}\left(\frac{\bar{W}^\theta}{\bar{L}}\right)^{\frac{1}{2(1-\theta\alpha)}},$$

$$L_i^R=x_{i,1}^{-\frac{1}{a^2-1}\frac{1}{1-\theta\beta}}x_{i,2}^{\frac{a}{a^2-1}\frac{1}{1-\theta\beta}}\bar{u}_i^{\frac{\theta}{1-\theta\beta}}\left(\frac{\bar{W}^\theta}{\bar{L}}\right)^{\frac{1}{2(1-\theta\beta)}}$$

or more succinctly as:

$$L_i^F=\bar{W}^{\rho_1}C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}} \tag{B.15}$$

$$L_i^R=\bar{W}^{\rho_2}C_{i,2}x_{i,1}^{b_{21}}x_{i,2}^{b_{22}} \tag{B.16}$$

where $C_{i,1}\equiv\frac{\bar{A}_i^{\frac{\theta}{1-\theta\alpha}}}{\bar{L}^{\frac{1}{2}\frac{1}{1-\theta\alpha}}}$,$C_{i,2}\equiv\frac{\bar{u}_i^{\frac{\theta}{1-\theta\beta}}}{\bar{L}^{\frac{1}{2}\frac{1}{1-\theta\beta}}}$, $b_{11}\equiv\frac{a}{a^2-1}\frac{1}{1-\theta\alpha}$, $b_{12}\equiv-\frac{1}{a^2-1}\frac{1}{1-\theta\alpha}$, $b_{21}\equiv-\frac{1}{a^2-1}\frac{1}{1-\theta\beta}$, $b_{22}\equiv\frac{a}{a^2-1}\frac{1}{1-\theta\beta}$, $\rho_1\equiv\frac{\theta}{2(1-\theta\alpha)}$ and $\rho_2\equiv\frac{\theta}{2(1-\theta\beta)}$. Substituting equations (B.15) and (B.16) into equations into the equilibrium system (B.8) and (B.9) yields:

$$x_{i,1}=W^{\rho_1}C_{i,1}x_{i,1}^{\frac{a}{a+1}+b_{11}}x_{i,2}^{-\frac{1}{a+1}+b_{12}}+\sum_j K_{ij}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}} \tag{B.17}$$

$$x_{i,2}=W^{\rho_2}C_{i,2}x_{i,1}^{-\frac{1}{a+1}+b_{21}}x_{i,2}^{\frac{a}{a+1}+b_{22}}+\sum_j K_{ji}x_{j,1}^{-\frac{a}{a+1}}x_{j,2}^{\frac{1}{a+1}}. \tag{B.18}$$

Equations (B.17) and (B.18) will form the basis of the uniqueness analysis below. For the analysis of existence, we proceed one step further and impose the normalization that $\sum_i L_i^F=\sum_i L_i^R=\bar{L}$,which then allows us to write:

$$L_i^F=\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}\bar{L}$$

$$L_i^R=\frac{C_{i,2}x_{i,1}^{b_{21}}x_{i,2}^{b_{22}}}{\sum_j C_{j,2}x_{j,1}^{b_{21}}x_{j,2}^{b_{22}}}\bar{L}$$

so that the system of equations becomes:

$$x_{i,1}=\bar{L}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}+\sum_j K_{ij}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}} \tag{B.19}$$

$$x_{i,2} = \bar{L} \frac{C_{i,2} x_{i,1}^{b_{21}} x_{i,2}^{b_{22}}}{\sum_j C_{j,2} x_{j,1}^{b_{21}} x_{j,2}^{b_{22}}} x_{i,1}^{-\frac{1}{a+1}} x_{i,2}^{\frac{a}{a+1}} + \sum_j K_{ji} x_{j,1}^{-\frac{a}{a+1}} x_{j,2}^{\frac{1}{a+1}}. \tag{B.20}$$

which is a special case of equations (B.1) and (B.2), as claimed.

B.2 Part 1 (Existence)

We first turn to the existence of the system. To do so, we rely on Brouwer's fixed point theorem. A natural way to proceed would be to define the operator $T(x) : \mathbb{R}_{++}^{2N} \to \mathbb{R}_{++}^{2N}$ such that:

$$T(x) = \begin{pmatrix} (T_1(x))_{i \in \mathcal{N}} \\ (T_2(x))_{i \in \mathcal{N}} \end{pmatrix} \equiv \begin{pmatrix} \left(D_{i,1} \frac{C_{i,1} x_{i,1}^{b_{11}} x_{i,2}^{b_{12}}}{\sum_{j \in \mathcal{N}} C_{j,1} x_{j,1}^{b_{11}} x_{j,2}^{b_{12}}} x_{i,1}^{\frac{a}{a+1}} x_{i,2}^{-\frac{1}{a+1}} + \sum_{j \in \mathcal{N}} K_{ij} x_{j,1}^{\frac{1}{a+1}} x_{j,2}^{-\frac{a}{a+1}} \right)_i \\ \left(D_{i,2} \frac{C_{i,2} x_{i,1}^{b_{21}} x_{i,2}^{b_{22}}}{\sum_{j \in \mathcal{N}} C_{j,2} x_{j,1}^{b_{21}} x_{j,2}^{b_{22}}} x_{i,1}^{-\frac{1}{a+1}} x_{i,2}^{\frac{a}{a+1}} + \sum_{j \in \mathcal{N}} K_{ji} x_{j,1}^{-\frac{a}{a+1}} x_{j,2}^{\frac{1}{a+1}} \right)_i \end{pmatrix}$$

While it is immediately obvious that T is continuous, unfortunately, T does not operate on a compact space, and hence Brouwer's theorem cannot be directly applied to it. Instead, following Allen, Arkolakis, and Takahashi (2020), we consider an alternative "scaled" system, the equilibrium of which will turn out to be a scaled version of the equilibrium of our system.

Define the operator: $\tilde{T}(x) \equiv \begin{pmatrix} \tilde{T}_1(x)_i \\ \tilde{T}_2(x)_i \end{pmatrix} : \mathbb{R}_{++}^{2N} \to \mathbb{R}_{++}^{2N}$ such that:

$$\tilde{T}(x) \equiv \begin{pmatrix} \left(\frac{D_{i,1} \frac{C_{i,1} x_{i,1}^{b_{11}} x_{i,2}^{b_{12}}}{\sum_j C_{j,1} x_{j,1}^{b_{11}} x_{j,2}^{b_{12}}} x_{i,1}^{\frac{a}{a+1}} x_{i,2}^{-\frac{1}{a+1}}}{\sum_i \left(D_{i,1} \frac{C_{i,1} x_{i,1}^{b_{11}} x_{i,2}^{b_{12}}}{\sum_j C_{j,1} x_{j,1}^{b_{11}} x_{j,2}^{b_{12}}} x_{i,1}^{\frac{a}{a+1}} x_{i,2}^{-\frac{1}{a+1}} + \sum_j K_{ij} x_{j,1}^{\frac{1}{a+1}} x_{j,2}^{-\frac{a}{a+1}} \right)} + \frac{\sum_j K_{ij} x_{j,1}^{\frac{1}{a+1}} x_{j,2}^{-\frac{a}{a+1}}}{\sum_i \left(D_{i,1} \frac{C_{i,1} x_{i,1}^{b_{11}} x_{i,2}^{b_{12}}}{\sum_j C_{j,1} x_{j,1}^{b_{11}} x_{j,2}^{b_{12}}} x_{i,1}^{\frac{a}{a+1}} x_{i,2}^{-\frac{1}{a+1}} + \sum_j K_{ij} x_{j,1}^{\frac{1}{a+1}} x_{j,2}^{-\frac{a}{a+1}} \right)} \right)_i \\ \left(\frac{D_{i,2} \frac{C_{i,2} x_{i,1}^{b_{21}} x_{i,2}^{b_{22}}}{\sum_j C_{j,2} x_{j,1}^{b_{21}} x_{j,2}^{b_{22}}} x_{i,1}^{-\frac{1}{a+1}} x_{i,2}^{\frac{a}{a+1}}}{\sum_i \left(D_{i,2} \frac{C_{i,2} x_{i,1}^{b_{21}} x_{i,2}^{b_{22}}}{\sum_j C_{j,2} x_{j,1}^{b_{21}} x_{j,2}^{b_{22}}} x_{i,1}^{-\frac{1}{a+1}} x_{i,2}^{\frac{a}{a+1}} + \sum_j K_{ji} x_{j,1}^{-\frac{a}{a+1}} x_{j,2}^{\frac{1}{a+1}} \right)} + \frac{\sum_j K_{ji} x_{j,1}^{-\frac{a}{a+1}} x_{j,2}^{\frac{1}{a+1}}}{\sum_i \left(D_{i,2} \frac{C_{i,2} x_{i,1}^{b_{21}} x_{i,2}^{b_{22}}}{\sum_j C_{j,2} x_{j,1}^{b_{21}} x_{j,2}^{b_{22}}} x_{i,1}^{-\frac{1}{a+1}} x_{i,2}^{\frac{a}{a+1}} + \sum_j K_{ji} x_{j,1}^{-\frac{a}{a+1}} x_{j,2}^{\frac{1}{a+1}} \right)} \right)_i \end{pmatrix},$$

where $x \equiv \begin{pmatrix} (x_{i,1})_i \\ (x_{i,2})_i \end{pmatrix} \in \mathbb{R}_{++}^{2N}$. Like with T, it is immediately obvious that $\tilde{T}$ is continuous. We show now that it also maps a compact space to itself, thereby allowing us to establish the existence of a fixed point of $\tilde{T}$.

Define the compact space $\mathcal{M} \subset \mathbb{R}_{++}^{2N} \equiv \left\{ x \in \mathbb{R}_{++}^{2N} | x_{il} \in [m_l, M_l] \ \forall l \in \{1,2\}, i \in \mathcal{N} \right\}$, where:

$$M_1 \equiv \max \left\{ \max_i \frac{D_{i,1}}{\sum_l D_{l,1}}, \max_{i,j} \frac{K_{ij}}{\sum_l K_{lj}} \right\}$$

$$m_1 \equiv \min \left\{ \min_i \frac{D_{i,1}}{\sum_l D_{l,1}}, \min_{i,j} \frac{K_{ij}}{\sum_l K_{lj}} \right\}$$

$$M_2 \equiv \max \left\{ \max_i \frac{D_{i,2}}{\sum_l D_{l,2}}, \max_{i,j} \frac{K_{ji}}{\sum_l K_{jl}} \right\}$$

$$m_2 \equiv \min \left\{ \min_i \frac{D_{i,1}}{\sum_l D_{l,1}}, \min_i \frac{K_{ij}}{\sum_l K_{lj}} \right\}$$

We claim that for all $x \in \mathcal{M}$, $\tilde{T}(x) \in \mathcal{M}$,i.e. $\tilde{T}$ operates from a compact space $\mathcal{M}$ to itself. We show the argument for $\tilde{T}_1(x) \leq M_1$; the argument for the remaining bounds proceed similarly. Consider the first term of the $\tilde{T}_1(x)$, $\tilde{T}_1^A(x) \equiv \frac{D_{i,1}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}}{\sum_i\left(D_{i,1}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}+\sum_j K_{ij}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}}\right)}$, which can be bounded as follows:

$$\begin{aligned}
\tilde{T}_1^A(x) &= \frac{D_{i,1}}{\sum_l D_{l,1}} \times \frac{\sum_l D_{l,1}}{D_{i,1}} \times \frac{D_{i,1}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}}{\sum_i\left(D_{i,1}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}+\sum_j K_{ij}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}}\right)} \iff \\
&= \frac{D_{i,1}}{\sum_l D_{l,1}} \times \frac{\sum_l D_{l,1}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}}{\sum_i\left(D_{i,1}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}+\sum_j K_{ij}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}}\right)} \implies \\
&\leq \left(\max_i \frac{D_{i,1}}{\sum_l D_{l,1}}\right) \times \frac{\sum_l D_{l,1}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}}{\sum_i\left(D_{i,1}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}+\sum_j K_{ij}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}}\right)}
\end{aligned}$$

Similarly, consider the second term of $\tilde{T}_1(x)$, $\tilde{T}_1^B(x) \equiv \frac{\sum_j K_{ij}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}}}{\sum_i\left(D_{i,1}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}+\sum_j K_{ij}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}}\right)}$,

which can be bounded as follows::

$$\begin{aligned}
\tilde{T}_1^B(x) &= \sum_j \frac{K_{ij}}{\sum_j K_{lj}} \times \frac{\sum_l K_{lj}}{K_{ij}} \times \frac{K_{ij}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}}}{\sum_i\left(D_{i,1}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}+\sum_j K_{ij}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}}\right)} \iff \\
&= \sum_j \frac{K_{ij}}{\sum_l K_{lj}} \times \frac{\sum_l \sum_j K_{lj}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}}}{\sum_i\left(D_{i,1}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}+\sum_j K_{ij}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}}\right)} \implies \\
&\leq \max_{i,j} \frac{K_{ij}}{\sum_l K_{lj}} \times \sum_j \frac{\left(\sum_l \sum_j K_{lj}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}}\right)}{\sum_i\left(D_{i,1}\frac{C_{i,1}x_{i,1}^{b_{11}}x_{i,2}^{b_{12}}}{\sum_j C_{j,1}x_{j,1}^{b_{11}}x_{j,2}^{b_{12}}}x_{i,1}^{\frac{a}{a+1}}x_{i,2}^{-\frac{1}{a+1}}+\sum_j K_{ij}x_{j,1}^{\frac{1}{a+1}}x_{j,2}^{-\frac{a}{a+1}}\right)}
\end{aligned}$$

so that together we have:

$$\tilde{T}_1(x)_i \leq \max\left\{\max_i \frac{D_{i,1}}{\sum_l D_{l,1}}, \max_{i,j} \frac{K_{ij}}{\sum_l K_{lj}}\right\},$$

as claimed.

Since $\tilde{T}$ is a continuous operator mapping a compact set $\mathcal{M}$ to itself, by Brouwer's fixed point, there exists a fixed point $\tilde{x}^* \in \mathcal{M}$ such that $\tilde{T}(\tilde{x}^*) = \tilde{x}^*$.

The next step of the existence proof is to show that there exists an equilibrium of the system characterized by equations (B.1) and (B.2), or, equivalently, there exists a fixed point x^* of the (un-scaled) operator T, i.e. $x^* = T(x^*)$. To do so, we will show that there exists a scalar $t > 0$ such that $x^* = t\tilde{x}^*$. Before we do this, however, we first need the following result that says the denominator of the $\tilde{T}_1(\tilde{x}^*)_i$ and $\tilde{T}_2(\tilde{x}^*)_i$ are equal, i.e.:

$$\sum_i \left(D_{i,1} \frac{C_{i,1} \left(\tilde{x}_{i,1}^*\right)^{b_{11}} \left(\tilde{x}_{i,2}^*\right)^{b_{12}}}{\sum_j C_{j,1} \left(\tilde{x}_{j,1}^*\right)^{b_{11}} \left(\tilde{x}_{j,2}^*\right)^{b_{12}}} \left(\tilde{x}_{i,1}^*\right)^{\frac{a}{a+1}} \left(\tilde{x}_{i,2}^*\right)^{-\frac{1}{a+1}} + \sum_j K_{ij} \left(\tilde{x}_{j,1}^*\right)^{\frac{1}{a+1}} \left(\tilde{x}_{j,2}^*\right)^{-\frac{a}{a+1}} \right) =$$

$$\sum_i \left(D_{i,2} \frac{C_{i,2} \left(\tilde{x}_{i,1}^*\right)^{b_{21}} \left(\tilde{x}_{i,2}^*\right)^{b_{22}}}{\sum_j C_{j,2} \left(\tilde{x}_{j,1}^*\right)^{b_{21}} \left(\tilde{x}_{j,2}^*\right)^{b_{22}}} \left(\tilde{x}_{i,1}^*\right)^{-\frac{1}{a+1}} \left(\tilde{x}_{i,2}^*\right)^{\frac{a}{a+1}} + \sum_j K_{ji} \left(\tilde{x}_{j,1}^*\right)^{-\frac{a}{a+1}} \left(\tilde{x}_{j,2}^*\right)^{\frac{1}{a+1}} \right) \tag{B.21}$$

The easiest way to see this is to re-write the scaled equilibrium conditions as functions of the endogenous variables $\{p_i, \pi_i\}$ (rather than $\{x_{i,1}, x_{i,2}\}$), which becomes:

$$p_i^a \pi_i = \frac{\delta_i p_i^{-(1-a)} + \sum_j K_{ij} \pi_j^{(1-a)}}{\sum_i \left(\delta_i p_i^{-(1-a)} + \sum_j K_{ij} \pi_j^{(1-a)} \right)}$$

$$p_i \pi_i^a = \frac{\gamma_i \pi_i^{-(1-a)} + \sum_j K_{ji} p_j^{(1-a)}}{\sum_i \left(\gamma_i \pi_i^{-(1-a)} + \sum_j K_{ji} p_j^{(1-a)} \right)}$$

Multiplying the first condition by p_i^{1-a} and summing over i yields:

$$\sum_i p_i \pi_i = \frac{\sum_i \delta_i + \sum_i \sum_j K_{ij} \pi_j^{(1-a)} p_i^{(1-a)}}{\sum_i \left(\delta_i p_i^{-(1-a)} + \sum_j K_{ij} \pi_j^{(1-a)} \right)}$$

Similarly, multiplying the second condition by π_i^{1-a} and summing over i yields:

$$\sum_i p_i \pi_i = \frac{\sum_i \gamma_i + \sum_i \sum_j K_{ji} p_j^{(1-a)} \pi_i^{(1-a)}}{\sum_i \left(\gamma_i \pi_i^{-(1-a)} + \sum_j K_{ji} p_j^{(1-a)} \right)}$$

Since the left hand side of both conditions are the same and the numerators on the right hand side are both the same (as recall $\sum_i \delta_i = \sum_i \gamma_i = \bar{L}$ in both the economic geography and urban models), the denominators must also be the same. Hence we have:

$$\sum_i \left(\delta_i p_i^{-(1-a)} + \sum_j K_{ij} \pi_j^{(1-a)} \right) = \sum_i \left(\gamma_i \pi_i^{-(1-a)} + \sum_j K_{ji} p_j^{(1-a)} \right),$$

or in $\{x_{i,1}, x_{i,2}\}$ space, equation (B.21) holds, as claimed.

Armed with this result, we are now prepared to construct a solution of equations (B.1) and (B.2). We posit that for all $i \in \mathcal{N}$ and $l \in \{1,2\}$ we have:

$$x_{i,l}^* = t \tilde{x}_{i,l}^* \tag{B.22}$$

where:

$$t^{\frac{2}{a+1}} = \sum_i \left(D_{i,1} \frac{C_{i,1} \left(\tilde{x}_{i,1}^*\right)^{b_{11}} \left(\tilde{x}_{i,2}^*\right)^{b_{12}}}{\sum_j C_{j,1} \left(\tilde{x}_{j,1}^*\right)^{b_{11}} \left(\tilde{x}_{j,2}^*\right)^{b_{12}}} \left(\tilde{x}_{i,1}^*\right)^{\frac{a}{a+1}} \left(\tilde{x}_{i,2}^*\right)^{-\frac{1}{a+1}} + \sum_j K_{ij} \left(\tilde{x}_{j,1}^*\right)^{\frac{1}{a+1}} \left(\tilde{x}_{j,2}^*\right)^{-\frac{a}{a+1}} \right) \iff \tag{B.23}$$

$$t^{\frac{2}{a+1}} = \sum_i \left(D_{i,2} \frac{C_{i,2} \left(\tilde{x}_{i,1}^*\right)^{b_{21}} \left(\tilde{x}_{i,2}^*\right)^{b_{22}}}{\sum_j C_{j,2} \left(\tilde{x}_{i,1}^*\right)^{b_{21}} \left(\tilde{x}_{i,2}^*\right)^{b_{22}}} \left(\tilde{x}_{i,1}^*\right)^{-\frac{1}{a+1}} \left(\tilde{x}_{i,2}^*\right)^{\frac{a}{a+1}} + \sum_j K_{ji} \left(\tilde{x}_{j,1}^*\right)^{-\frac{a}{a+1}} \left(\tilde{x}_{j,2}^*\right)^{\frac{1}{a+1}} \right), \tag{B.24}$$

where the second line immediately follows from equation (B.21).

Since $\tilde{x}^* = \tilde{T}(\tilde{x}^*)$, we first consider $\tilde{x}^*_1 = \tilde{T}_1(\tilde{x}^*)$. Imposing (B.22) and (B.23) yields:

$$\begin{aligned}
\frac{1}{t}x^*_{i,1} =& \frac{D_{i,1}\frac{C_{i,1}\left(\frac{1}{t}x^*_{i,1}\right)^{b_{11}}\left(\frac{1}{t}x^*_{i,2}\right)^{b_{12}}}{\sum_j C_{j,1}\left(\frac{1}{t}x^*_{j,1}\right)^{b_{11}}\left(\frac{1}{t}x^*_{j,2}\right)^{b_{12}}}\left(\frac{1}{t}x^*_{i,1}\right)^{\frac{a}{a+1}}\left(\frac{1}{t}x^*_{i,2}\right)^{-\frac{1}{a+1}}}{\sum_i\left(D_{i,1}\frac{C_{i,1}\left(\tilde{x}^*_{i,1}\right)^{b_{11}}\left(\tilde{x}^*_{i,2}\right)^{b_{12}}}{\sum_j C_{j,1}\left(\tilde{x}^*_{j,1}\right)^{b_{11}}\left(\tilde{x}^*_{j,2}\right)^{b_{12}}}\left(\tilde{x}^*_{i,1}\right)^{\frac{a}{a+1}}\left(\tilde{x}^*_{i,2}\right)^{-\frac{1}{a+1}}+\sum_j K_{ij}\left(\tilde{x}^*_{j,1}\right)^{\frac{1}{a+1}}\left(\tilde{x}^*_{j,2}\right)^{-\frac{a}{a+1}}\right)}+ \\
& \frac{\sum_j K_{ij}\left(\frac{1}{t}x^*_{j,1}\right)^{\frac{1}{a+1}}\left(\frac{1}{t}x^*_{j,2}\right)^{-\frac{a}{a+1}}}{\sum_i\left(D_{i,1}\frac{C_{i,1}\left(\tilde{x}^*_{i,1}\right)^{b_{11}}\left(\tilde{x}^*_{i,2}\right)^{b_{12}}}{\sum_j C_{j,1}\left(\tilde{x}^*_{j,1}\right)^{b_{11}}\left(\tilde{x}^*_{j,2}\right)^{b_{12}}}\left(\tilde{x}^*_{i,1}\right)^{\frac{a}{a+1}}\left(\tilde{x}^*_{i,2}\right)^{-\frac{1}{a+1}}+\sum_j K_{ij}\left(\tilde{x}^*_{j,1}\right)^{\frac{1}{a+1}}\left(\tilde{x}^*_{j,2}\right)^{-\frac{a}{a+1}}\right)} \iff \\
\left(\frac{1}{t}\right)^{\frac{2}{a+1}} x^*_{i,1} =& \frac{D_{i,1}\frac{C_{i,1}\left(x^*_{i,1}\right)^{b_{11}}\left(x^*_{i,2}\right)^{b_{12}}}{\sum_j C_{j,1}\left(x^*_{j,1}\right)^{b_{11}}\left(x^*_{j,2}\right)^{b_{12}}}\left(x^*_{i,1}\right)^{\frac{a}{a+1}}\left(x^*_{i,2}\right)^{-\frac{1}{a+1}}}{\sum_i\left(D_{i,1}\frac{C_{i,1}\left(\tilde{x}^*_{i,1}\right)^{b_{11}}\left(\tilde{x}^*_{i,2}\right)^{b_{12}}}{\sum_j C_{j,1}\left(\tilde{x}^*_{j,1}\right)^{b_{11}}\left(\tilde{x}^*_{j,2}\right)^{b_{12}}}\left(\tilde{x}^*_{i,1}\right)^{\frac{a}{a+1}}\left(\tilde{x}^*_{i,2}\right)^{-\frac{1}{a+1}}+\sum_j K_{ij}\left(\tilde{x}^*_{j,1}\right)^{\frac{1}{a+1}}\left(\tilde{x}^*_{j,2}\right)^{-\frac{a}{a+1}}\right)} \\
& + \frac{\sum_j K_{ij}\left(x^*_{j,1}\right)^{\frac{1}{a+1}}\left(x^*_{j,2}\right)^{-\frac{a}{a+1}}}{\sum_i\left(D_{i,1}\frac{C_{i,1}\left(\tilde{x}^*_{i,1}\right)^{b_{11}}\left(\tilde{x}^*_{i,2}\right)^{b_{12}}}{\sum_j C_{j,1}\left(\tilde{x}^*_{j,1}\right)^{b_{11}}\left(\tilde{x}^*_{j,2}\right)^{b_{12}}}\left(\tilde{x}^*_{i,1}\right)^{\frac{a}{a+1}}\left(\tilde{x}^*_{i,2}\right)^{-\frac{1}{a+1}}+\sum_j K_{ij}\left(\tilde{x}^*_{j,1}\right)^{\frac{1}{a+1}}\left(\tilde{x}^*_{j,2}\right)^{-\frac{a}{a+1}}\right)} \iff \\
x^*_{i,1} =& D_{i,1}\frac{C_{i,1}\left(x^*_{i,1}\right)^{b_{11}}\left(x^*_{i,2}\right)^{b_{12}}}{\sum_j C_{j,1}\left(x^*_{j,1}\right)^{b_{11}}\left(x^*_{j,2}\right)^{b_{12}}}\left(x^*_{i,1}\right)^{\frac{a}{a+1}}\left(x^*_{i,2}\right)^{-\frac{1}{a+1}}+\sum_j K_{ij}\left(x^*_{j,1}\right)^{\frac{1}{a+1}}\left(x^*_{j,2}\right)^{-\frac{a}{a+1}},
\end{aligned}$$

i.e. $x^*_1 = \tilde{T}_1(x^*)$ holds as claimed.

Similarly, considering $\tilde{x}^*_2 = \tilde{T}_2(\tilde{x}^*)$ and imposing (B.22) and (B.24) yields:

$$\begin{aligned}
\frac{1}{t}x^*_{i,2} =& \frac{D_{i,2}\frac{C_{i,2}\left(\frac{1}{t}x^*_{i,1}\right)^{b_{21}}\left(\frac{1}{t}x^*_{i,2}\right)^{b_{22}}}{\sum_j C_{j,2}\left(\frac{1}{t}x^*_{j,1}\right)^{b_{21}}\left(\frac{1}{t}x^*_{j,2}\right)^{b_{22}}}\left(\frac{1}{t}x^*_{i,1}\right)^{-\frac{1}{a+1}}\left(\frac{1}{t}x^*_{i,2}\right)^{\frac{a}{a+1}}}{\sum_i\left(D_{i,2}\frac{C_{i,2}\left(\tilde{x}^*_{i,1}\right)^{b_{21}}\left(\tilde{x}^*_{i,2}\right)^{b_{22}}}{\sum_j C_{j,2}\left(\tilde{x}^*_{i,1}\right)^{b_{21}}\left(\tilde{x}^*_{i,2}\right)^{b_{22}}}\left(\tilde{x}^*_{i,1}\right)^{-\frac{1}{a+1}}\left(\tilde{x}^*_{i,2}\right)^{\frac{a}{a+1}}+\sum_j K_{ji}\left(\tilde{x}^*_{j,1}\right)^{-\frac{a}{a+1}}\left(\tilde{x}^*_{j,2}\right)^{\frac{1}{a+1}}\right)}+ \\
& \frac{\sum_j K_{ji}\left(\frac{1}{t}x^*_{j,1}\right)^{-\frac{a}{a+1}}\left(\frac{1}{t}x^*_{j,2}\right)^{\frac{1}{a+1}}}{\sum_i\left(D_{i,2}\frac{C_{i,2}\left(\tilde{x}^*_{i,1}\right)^{b_{21}}\left(\tilde{x}^*_{i,2}\right)^{b_{22}}}{\sum_j C_{j,2}\left(\tilde{x}^*_{i,1}\right)^{b_{21}}\left(\tilde{x}^*_{i,2}\right)^{b_{22}}}\left(\tilde{x}^*_{i,1}\right)^{-\frac{1}{a+1}}\left(\tilde{x}^*_{i,2}\right)^{\frac{a}{a+1}}+\sum_j K_{ji}\left(\tilde{x}^*_{j,1}\right)^{-\frac{a}{a+1}}\left(\tilde{x}^*_{j,2}\right)^{\frac{1}{a+1}}\right)} \iff \\
\left(\frac{1}{t}\right)^{\frac{2}{a+1}} x^*_{i,2} =& \frac{D_{i,2}\frac{C_{i,2}\left(x^*_{i,1}\right)^{b_{21}}\left(x^*_{i,2}\right)^{b_{22}}}{\sum_j C_{j,2}\left(x^*_{j,1}\right)^{b_{21}}\left(x^*_{j,2}\right)^{b_{22}}}\left(x^*_{i,1}\right)^{-\frac{1}{a+1}}\left(x^*_{i,2}\right)^{\frac{a}{a+1}}}{\sum_i\left(D_{i,2}\frac{C_{i,2}\left(\tilde{x}^*_{i,1}\right)^{b_{21}}\left(\tilde{x}^*_{i,2}\right)^{b_{22}}}{\sum_j C_{j,2}\left(\tilde{x}^*_{i,1}\right)^{b_{21}}\left(\tilde{x}^*_{i,2}\right)^{b_{22}}}\left(\tilde{x}^*_{i,1}\right)^{-\frac{1}{a+1}}\left(\tilde{x}^*_{i,2}\right)^{\frac{a}{a+1}}+\sum_j K_{ji}\left(\tilde{x}^*_{j,1}\right)^{-\frac{a}{a+1}}\left(\tilde{x}^*_{j,2}\right)^{\frac{1}{a+1}}\right)}+ \\
& \frac{\sum_j K_{ji}\left(x^*_{j,1}\right)^{-\frac{a}{a+1}}\left(x^*_{j,2}\right)^{\frac{1}{a+1}}}{\sum_i\left(D_{i,2}\frac{C_{i,2}\left(\tilde{x}^*_{i,1}\right)^{b_{21}}\left(\tilde{x}^*_{i,2}\right)^{b_{22}}}{\sum_j C_{j,2}\left(\tilde{x}^*_{i,1}\right)^{b_{21}}\left(\tilde{x}^*_{i,2}\right)^{b_{22}}}\left(\tilde{x}^*_{i,1}\right)^{-\frac{1}{a+1}}\left(\tilde{x}^*_{i,2}\right)^{\frac{a}{a+1}}+\sum_j K_{ji}\left(\tilde{x}^*_{j,1}\right)^{-\frac{a}{a+1}}\left(\tilde{x}^*_{j,2}\right)^{\frac{1}{a+1}}\right)} \iff \\
x^*_{i,2} =& D_{i,2}\frac{C_{i,2}\left(x^*_{i,1}\right)^{b_{21}}\left(x^*_{i,2}\right)^{b_{22}}}{\sum_j C_{j,2}\left(x^*_{j,1}\right)^{b_{21}}\left(x^*_{j,2}\right)^{b_{22}}}\left(x^*_{i,1}\right)^{-\frac{1}{a+1}}\left(x^*_{i,2}\right)^{\frac{a}{a+1}}+\sum_j K_{ji}\left(x^*_{j,1}\right)^{-\frac{a}{a+1}}\left(x^*_{j,2}\right)^{\frac{1}{a+1}},
\end{aligned}$$

i.e. $x^*_2 = \tilde{T}_2(x^*)$ as well. Together, this means we have successfully constructed a fixed point x^* of the (un-scaled) operator T, i.e. $x^* = T(x^*)$, i.e. there exists a solution to the system characterized by equations (B.1) and (B.2).

Finally, we show that the equilibrium we have found is strictly positive, i.e. $x_{i,1} > 0$ for all $i \in \mathcal{N}$ and $l \in \{1, 2\}$. (Note that all $x^*_{i,l}$ are trivially finite since $t \in (0, \infty)$ and $\tilde{x}^* \in \mathcal{M}$, so that $x^*_{i,l} \leq tM_l < \infty$). We proceed by contradiction: suppose not. Then there exists an $i \in \mathcal{N}$ and $l \in \{1, 2\}$ such that $x^*_{i,l} = 0$.

Suppose that $l=1$ (the case of $l=2$ proceeds analogously). We then have:

$$x_{i,1}^{*}=D_{i,1}\frac{C_{i,1}\left(x_{i,1}^{*}\right)^{b_{11}}\left(x_{i,2}^{*}\right)^{b_{12}}}{\sum_{j}C_{j,1}\left(x_{j,1}^{*}\right)^{b_{11}}\left(x_{j,2}^{*}\right)^{b_{12}}}\left(x_{i,1}^{*}\right)^{\frac{a}{a+1}}\left(x_{i,2}^{*}\right)^{-\frac{1}{a+1}}+\sum_{j}K_{ij}\left(x_{j,1}^{*}\right)^{\frac{1}{a+1}}\left(x_{j,2}^{*}\right)^{-\frac{a}{a+1}}=0$$

Since $\left\{x_{j,1}^{*}\right\}$ are weakly positive and finite, since $K_{ij}\geq 0$, and $a\in[0,1]$, for this to be true, this means that for all j connected to i (i.e. the set of j such that $K_{ij}>0$), it must be the case that $x_{j,1}^{*}=0$ as well. The same argument implies that all j' connected to any of these j also have $x_{j',1}^{*}=0$. Since the matrix $\mathbf{K}=[K_{ij}]$ is connected, repeating this argument iteratively then implies that $x_{j,1}^{*}=0$ for all j. But if $x_{j,1}^{*}=0$ for all j, then from equation (B.2) $x_{j,2}^{*}$ would be infinite for all j, which is a contradiction, as $x_{i,2}^{*}\leq tM_2$. Hence, the equilibrium is strictly positive.

B.3 Part 2 (Uniqueness)

We now proceed to study the uniqueness of the system. We consider first an economic geography with a symmetric infrastructure matrix; we then consider an urban model.

B.3.1 Part 2(a): Uniqueness in an Economic Geography Model with Symmetric Infrastructure Matrix

Consider an economic geography model with a symmetric infrastructure matrix, i.e. $\bar{t}_{kl}=\bar{t}_{lk}$ for all $l,k\in\mathcal{N}\times\mathcal{N}$. It is well known (see e.g. Anderson and Van Wincoop (2003); Allen, Arkolakis, and Takahashi (2020)) that with symmetric transportation costs, the market access terms are equal up to scale). It turns out that this is also true about a symmetric infrastructure matrix in the presence of endogenous traffic congestion. To see this, note a symmetric infrastructure matrix implies $K_{ij}=K_{ji}$ for all i and j, so that equations (B.6) and (B.7) can be written in the economic geography case as:

$$p_i^{a}\pi_i=Y_i p_i^{-(1-a)}+\sum_j K_{ij}\pi_j^{(1-a)} \tag{B.25}$$

$$p_i\pi_i^{a}=Y_i\pi_i^{-(1-a)}+\sum_j K_{ij}p_j^{(1-a)}. \tag{B.26}$$

It is obvious by inspection of equations (B.25) and (B.26) that a solution to this system is $\pi_i=\kappa p_i$ for some $\kappa>0$. (Indeed, one can show using the same tools applied to the urban model in the next section that this is the unique solution). As a result, it is sufficient to focus on the solution of the single equation case, which can be written as follows:

$$\kappa^{a}p_i^{1+a}=\kappa^{a-1}Y_i p_i^{-(1-a)}+\sum_j K_{ij}p_j^{(1-a)}$$

Recall from Section B.1.1 that we can write:

$$Y_i=p_i^{\frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}}\pi_i^{\frac{(1-\beta)}{1-\beta-\theta(\alpha+\beta)}}\left(\frac{\bar{u}_i}{W}\right)^{\theta\frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}}\bar{A}_i^{\theta\frac{(1-\beta)}{1-\beta-\theta(\alpha+\beta)}}.$$

so that the full system becomes:

$$\kappa^{a}p_i^{1+a}=\kappa^{(a-1)}\bar{W}^{-\theta\frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}}\bar{u}_i^{\theta\frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}}\bar{A}_i^{\theta\frac{(1-\beta)}{1-\beta-\theta(\alpha+\beta)}}p_i^{\frac{(1+\alpha)+(1-\beta)}{1-\beta-\theta(\alpha+\beta)}-(1-a)}p_i^{-(1-a)}+\sum_j K_{ij}p_j^{(1-a)}.$$

Define $x_i\equiv\kappa^{a}p_i^{1+a}$ so that $p_i=\kappa^{-\frac{a}{1+a}}x_i^{\frac{1}{1+a}}$ so that we can write:

$$x_i=\kappa^{-\frac{a}{1+a}\frac{(1+\alpha)+(1-\beta)}{1-\beta-\theta(\alpha+\beta)}-\left(\frac{1-a}{1+a}\right)}\bar{W}^{-\theta\frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}}\bar{u}_i^{\theta\frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}}\bar{A}_i^{\theta\frac{(1-\beta)}{1-\beta-\theta(\alpha+\beta)}}x_i^{\frac{1}{1+a}\frac{(1+\alpha)+(1-\beta)}{1-\beta-\theta(\alpha+\beta)}}x_i^{-\left(\frac{1-a}{1+a}\right)}+\kappa^{-\frac{a(1-a)}{1+a}}\sum_j K_{ij}x_j^{\frac{1-a}{1+a}},$$

or, written slightly more succinctly:

$$x_i = \kappa^{\rho_1} \bar{W}^{\rho_2} C_i x_i^{b-\left(\frac{1-a}{1+a}\right)} + \kappa^{\rho_3} \sum_j K_{ij} x_j^{\frac{1-a}{1+a}}, \tag{B.27}$$

where $\rho_1 \equiv -\frac{a}{1+a} \frac{(1+\alpha)+(1-\beta)}{1-\beta-\theta(\alpha+\beta)} - \left(\frac{1-a}{1+a}\right)$, $\rho_2 \equiv -\theta \frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}$, $\rho_3 \equiv -\frac{a(1-a)}{1+a}$,$b \equiv \frac{1}{1+a} \frac{(1+\alpha)+(1-\beta)}{1-\beta-\theta(\alpha+\beta)}$, and $C_i \equiv \bar{u}_i^{\theta \frac{(1+\alpha)}{1-\beta-\theta(\alpha+\beta)}} \bar{A}_i^{\theta \frac{(1-\beta)}{1-\beta-\theta(\alpha+\beta)}}$.

We proceed by contradiction. Suppose there are two solutions $\left\{W_y, \kappa_y, \{y_i\}_{i\in\mathcal{N}}\right\}$ and $\left\{W_x, \kappa_x, \{x_i\}_{i\in\mathcal{N}}\right\}$ that both solve the system of equations (B.27) and that the two equations are distinct, i.e. they are not equal up to scale. To proceed, we take ratios of the two assumed solutions:

$$\begin{aligned} \frac{y_i}{x_i} &= \frac{\kappa_y^{\rho_1} \bar{W_y}^{\rho_2} C_i y_i^{b-\left(\frac{1-a}{1+a}\right)} + \kappa_y^{\rho_3} \sum_j K_{ij} y_j^{\frac{1-a}{1+a}}}{\kappa_x^{\rho_1} \bar{W_x}^{\rho_2} C_i x_i^{b-\left(\frac{1-a}{1+a}\right)} + \kappa_x^{\rho_3} \sum_j K_{ij} x_j^{\frac{1-a}{1+a}}} \iff \\ \hat{x}_i &= D_i \hat{\kappa}^{\rho_1} \hat{W}^{\rho_2} \hat{x}_i^{b-\left(\frac{1-a}{1+a}\right)} + \sum_j F_{ij} \hat{\kappa}^{\rho_3} \hat{x}_j^{\frac{1-a}{1+a}}, \end{aligned} \tag{B.28}$$

where $\hat{x}_i \equiv \frac{y_i}{x_i}$,$\hat{\kappa} \equiv \frac{\kappa_y}{\kappa_x}$, $\hat{W} \equiv \frac{W_y}{W_x}$, $D_i \equiv \frac{\kappa_x^{\rho_1} \bar{W_x}^{\rho_2} C_i x_i^{b-\left(\frac{1-a}{1+a}\right)}}{\kappa_x^{\rho_1} \bar{W_x}^{\rho_2} C_i x_i^{b-\left(\frac{1-a}{1+a}\right)} + \kappa_x^{\rho_3} \sum_j K_{ij} x_j^{\frac{1-a}{1+a}}}$, and $F_{ij} \equiv \frac{\kappa_x^{\rho_3} K_{ij} x_j^{\frac{1-a}{1+a}}}{\kappa_x^{\rho_1} \bar{W_x}^{\rho_2} C_i x_i^{b-\left(\frac{1-a}{1+a}\right)} + \kappa_x^{\rho_3} \sum_j K_{ij} x_j^{\frac{1-a}{1+a}}}$.

We now construct a maximum and minimum bound to equation (B.28). Define $M \equiv \max_i \hat{x}_i$, $m \equiv \min_i \hat{x}_i$, and $\mu \equiv M/m$. Note that because $a \in [0,1]$, we have:

$$\hat{x}_i \leq D_i \hat{\kappa}^{\rho_1} \hat{W}^{\rho_2} \hat{x}_i^{b-\left(\frac{1-a}{1+a}\right)} + \sum_j F_{ij} \hat{\kappa}^{\rho_3} \max_j x_j^{\frac{1-a}{1+a}}. \tag{B.29}$$

Indeed (and importantly), the inequality is strict. How do you see this? Well, suppose the inequality is not strict. Then there exists an $i \in S$ such that:

$$\hat{x}_{i,1} = D_i \hat{\kappa}^{\rho_1} \hat{W}^{\rho_2} \hat{x}_i^{b-\left(\frac{1-a}{1+a}\right)} + \sum_j F_{ij} \hat{\kappa}^{\rho_3} \max_j x_j^{\frac{1-a}{1+a}}.$$

For this to be true, it must be the case that for all j such that $F_{ij} > 0$ that $\hat{x}_j^{\frac{1}{a+1}} = \max_i \hat{x}_j^{\frac{1}{a+1}}$, or equivalently, $\hat{x}_j = \max_j \hat{x}_j = \hat{x}_i$. Since all the locations are connected, we can choose any of these j as our new i, they must also satisfy the equation with equality, and the argument can continue to the point that we have $\hat{x}_i = \hat{x}_j$ for all i and j. But this is a contradiction, since we assumed that there are two distinct solutions. Hence the inequality is strict. As a result, equation (B.29) implies:

$$M < D_i \hat{\kappa}^{\rho_1} \hat{W}^{\rho_2} M^{1^+\left(b-\left(\frac{1-a}{1+a}\right)\right)} m^{1^-\left(b-\left(\frac{1-a}{1+a}\right)\right)} + \sum_j F_{ij} \hat{\kappa}^{\rho_3} M^{\frac{1-a}{a+1}}, \tag{B.30}$$

where the (apologetically cumbersome) notation $1^+(x) \equiv \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$ and $1^-(x) \equiv \begin{cases} x & \text{if } x < 0 \\ 0 & \text{if } x \geq 0 \end{cases}$, is necessary to consider multiple cases of the signs of the exponents at once. A similar argument can be made to establish the following lower bound, resulting in:

$$m > D_i \hat{\kappa}^{\rho_1} \hat{W}^{\rho_2} m^{1^+\left(b-\left(\frac{1-a}{1+a}\right)\right)} M^{1^-\left(b-\left(\frac{1-a}{1+a}\right)\right)} + \sum_j F_{ij} \hat{\kappa}^{\rho_2} m^{\frac{1-a}{a+1}}. \tag{B.31}$$

Taking ratios of the upper to lower bounds achieves a strict upper bound on μ:

$$\mu < \frac{D_i \hat{\kappa}^{\rho_1} \hat{W}^{\rho_2} M^{1^+\left(b-\left(\frac{1-a}{1+a}\right)\right)} m^{1^-\left(b-\left(\frac{1-a}{1+a}\right)\right)} + \sum_j F_{ij} \hat{\kappa}^{\rho_3} M^{\frac{1-a}{a+1}}}{D_i \hat{\kappa}^{\rho_1} \hat{W}^{\rho_2} m^{1^+\left(b-\left(\frac{1-a}{1+a}\right)\right)} M^{1^-\left(b-\left(\frac{1-a}{1+a}\right)\right)} + \sum_j F_{ij} \hat{\kappa}^{\rho_2} m^{\frac{1-a}{a+1}}} \iff$$

$$\mu < \beta \mu^{\left|b-\left(\frac{1-a}{1+a}\right)\right|} + (1-\beta)\, \mu^{\frac{1-a}{a+1}}, \tag{B.32}$$

where $\beta \equiv \frac{D_i \hat{\kappa}^{\rho_1} \hat{W}^{\rho_2} m^{1^+\left(b-\left(\frac{1-a}{1+a}\right)\right)} M^{1^-\left(b-\left(\frac{1-a}{1+a}\right)\right)}}{D_i \hat{\kappa}^{\rho_1} \hat{W}^{\rho_2} m^{1^+\left(b-\left(\frac{1-a}{1+a}\right)\right)} M^{1^-\left(b-\left(\frac{1-a}{1+a}\right)\right)} + \sum_j F_{ij} \hat{\kappa}^{\rho_2} m^{\frac{1-a}{a+1}}}$. Equation (B.32) says that μ is bounded above by a weighted average of two terms. For this to be true, it must either be the case that μ is smaller than the first of the two terms or μ is smaller than the second of the two terms (or both). Since $\mu \geq 1$, we then have a contradiction if both

$$\left|\frac{1-a}{a+1}\right| \leq 1,$$

which is assured since $a \in [0,1]$ and:

$$\left|b - \left(\frac{1-a}{1+a}\right)\right| \leq 1 \tag{B.33}$$

Hence a contradiction arises (and therefore uniqueness is assured) if equation (B.33) holds. Recall from above that $a \equiv \frac{\theta\lambda}{1+\theta\lambda}$ and $b \equiv \frac{1}{1+a} \frac{(1+\alpha)+(1-\beta)}{1-\beta-\theta(\alpha+\beta)}$. It is straightforward (but tedious) to verify that for all $\alpha \in [-1,1]$, $\beta \in [-1,1]$, $\lambda \geq 0$, and $\theta \geq 0$ the following condition ensures equation (B.33) is satisfied an uniqueness is assured:

$$\alpha + \beta \leq 0,$$

as claimed.

B.3.2 Part 2(b): Uniqueness in an Urban Model

From equations (B.17) and (B.18), we have that the equilibrium of the urban model satisfies the following system of equations:

$$x_{i,1} = W^{\rho_1} C_{i,1} x_{i,1}^{\frac{a}{a+1}+b_{11}} x_{i,2}^{-\frac{1}{a+1}+b_{12}} + \sum_j K_{ij} x_{j,1}^{\frac{1}{a+1}} x_{j,2}^{-\frac{a}{a+1}} \tag{B.34}$$

$$x_{i,2} = W^{\rho_2} C_{i,2} x_{i,1}^{-\frac{1}{a+1}+b_{21}} x_{i,2}^{\frac{a}{a+1}+b_{22}} + \sum_j K_{ji} x_{j,1}^{-\frac{a}{a+1}} x_{j,2}^{\frac{1}{a+1}}. \tag{B.35}$$

As in the economic geography model, we prove uniqueness by contradiction. Suppose there exists a $\left\{W_y, \{y_{i,1}\}_i, \{y_{i,2}\}_i\right\}$ and $\left\{W_x, \{x_{i,1}\}_i, \{x_{i,2}\}_i\right\}$ that both solve equations (B.34) and (B.35), where solutions $\{y_{i,1}, y_{i,2}\}$ and $\{x_{i,1}, x_{i,2}\}$ are not equal up to scale. We will derive conditions under which this implies a contradiction.

The first step is to re-write the system in terms of ratios of the two proposed solutions. Beginning with equation (B.34) , we have:

$$\frac{y_{i,1}}{x_{i,1}} = \frac{W_y^{\rho_1} C_{i,1} y_{i,1}^{\frac{a}{a+1}+b_{11}} y_{i,2}^{-\frac{1}{a+1}+b_{12}} + \sum_j K_{ij} y_{j,1}^{\frac{1}{a+1}} y_{j,2}^{-\frac{a}{a+1}}}{W_x^{\rho_1} C_{i,1} x_{i,1}^{\frac{a}{a+1}+b_{11}} x_{i,2}^{-\frac{1}{a+1}+b_{12}} + \sum_j K_{ij} x_{j,1}^{\frac{1}{a+1}} x_{j,2}^{-\frac{a}{a+1}}} \iff$$

$$\hat{x}_{i,1} = D_{i,1} \hat{W}^{\rho_1} \left(\hat{x}_{i,1}\right)^{\frac{a}{a+1}+b_{11}} \left(\hat{x}_{i,2}\right)^{-\frac{1}{a+1}+b_{12}} + \sum_j F_{ij,1} \left(\hat{x}_{j,1}\right)^{\frac{1}{a+1}} \left(\hat{x}_{j,2}\right)^{-\frac{a}{a+1}}, \tag{B.36}$$

where $\hat{x}_{i,1} \equiv \frac{y_{i,1}}{x_{i,1}}$, $\hat{W} \equiv \frac{W_y}{W_x}$,

$$D_{i,1} \equiv \frac{W_x^{\rho_1} C_{i,1} x_{i,1}^{\frac{a}{a+1}+b_{11}} x_{i,2}^{-\frac{1}{a+1}+b_{12}}}{W_x^{\rho_1} C_{i,1} x_{i,1}^{\frac{a}{a+1}+b_{11}} x_{i,2}^{-\frac{1}{a+1}+b_{12}} + \sum_j K_{ij} x_{j,1}^{\frac{1}{a+1}} x_{j,2}^{-\frac{a}{a+1}}},$$

$$F_{ij,1} \equiv \frac{K_{ij}}{W_x^{\rho_1} C_{i,1} x_{i,1}^{\frac{a}{a+1}+b_{11}} x_{i,2}^{-\frac{1}{a+1}+b_{12}} + \sum_j K_{ij} x_{j,1}^{\frac{1}{a+1}} x_{j,2}^{-\frac{a}{a+1}}}.$$

Note that $D_{i,1} + \sum_j F_{ij,1} = 1$ for all $i \in \mathcal{N}$. Similarly, equation (B.35) can be written in ratios as follows:

$$\hat{x}_{i,2} = D_{i,2} \hat{W}^{\rho_2} \left(\hat{x}_{i,1}\right)^{-\frac{1}{a+1}+b_{21}} \left(\hat{x}_{i,2}\right)^{\frac{a}{a+1}+b_{22}} + \sum_j F_{ij,2} \left(\hat{x}_{j,1}\right)^{-\frac{a}{a+1}} \left(\hat{x}_{j,2}\right)^{\frac{1}{a+1}}, \tag{B.37}$$

where $\hat{x}_{i,2} \equiv \frac{y_{i,2}}{x_{i,2}}$,

$$D_{i,2} \equiv \frac{W^{\rho_2} C_{i,2} x_{i,1}^{-\frac{1}{a+1}+b_{21}} x_{i,2}^{\frac{a}{a+1}+b_{22}}}{W^{\rho_2} C_{i,2} x_{i,1}^{-\frac{1}{a+1}+b_{21}} x_{i,2}^{\frac{a}{a+1}+b_{22}} + \sum_j K_{ji} x_{j,1}^{-\frac{a}{a+1}} x_{j,2}^{\frac{1}{a+1}}},$$

$$F_{ij,2} \equiv \frac{K_{ji}}{W^{\rho_2} C_{i,2} x_{i,1}^{-\frac{1}{a+1}+b_{21}} x_{i,2}^{\frac{a}{a+1}+b_{22}} + \sum_j K_{ji} x_{j,1}^{-\frac{a}{a+1}} x_{j,2}^{\frac{1}{a+1}}}.$$

Note that $D_{i,2} + \sum_j F_{ij,2} = 1$ for all $i \in \mathcal{N}$.

In what follows, we focus on the system of equations (B.36) and (B.37) written in ratios. Because we are assuming the solutions $\{y_{i,1}, y_{i,2}\}$ and $\{x_{i,1}, x_{i,2}\}$ are not equal up to scale, it must be the case that there is at least one $\hat{x}_{i,l} \neq 1$.

To proceed, we start bounding the ratios of the solutions. Define $M_l \equiv \max_i \hat{x}_{i,l}$, $m_l \equiv \min_i \hat{x}_{i,l}$, and $\mu_l \equiv \frac{M_l}{m_l}$. From equation (B.36) we have for all $i \in \mathcal{N}$:

$$\hat{x}_{i,1} \leq D_{i,1} \hat{W}^{\rho_1} \left(\hat{x}_{i,1}\right)^{\frac{a}{a+1}+b_{11}} \left(\hat{x}_{i,2}\right)^{-\frac{1}{a+1}+b_{12}} + \sum_j F_{ij,1} \max_i \left(\left(\hat{x}_{j,1}\right)^{\frac{1}{a+1}}\right) \max_j \left(\left(\hat{x}_{j,2}\right)^{-\frac{a}{a+1}}\right). \tag{B.38}$$

Indeed (and importantly), the inequality is strict. How do you see this? Well, suppose the inequality is not strict. Then there exists an $i \in S$ such that:

$$\hat{x}_{i,1} = D_{i,1} \hat{W}^{\rho_1} \left(\hat{x}_{i,1}\right)^{\frac{a}{a+1}+b_{11}} \left(\hat{x}_{i,2}\right)^{-\frac{1}{a+1}+b_{12}} + \sum_j F_{ij,1} \max_i \left(\left(\hat{x}_{j,1}\right)^{\frac{1}{a+1}}\right) \max_j \left(\left(\hat{x}_{j,2}\right)^{-\frac{a}{a+1}}\right).$$

For this to be true, it must be the case that for all j such that $F_{ij,1} > 0$ that $\left(\hat{x}_{j,1}\right)^{\frac{1}{a+1}} = \max_i \left(\left(\hat{x}_{j,1}\right)^{\frac{1}{a+1}}\right)$, or equivalently, $\hat{x}_{j,1} = \max_j \hat{x}_{j,1} = \hat{x}_{i,1}$. Similarly, it must be the case that $\hat{x}_{j,2} = \min_j \hat{x}_{j,2}$. Since all the locations are connected, we can choose any of these j as our new i, they must also satisfy the equation with equality, and the argument can continue to the point that we have $\hat{x}_{i,1} = \hat{x}_{j,1}$ and $\hat{x}_{i,2} = \hat{x}_{j,2}$ for all i and j. But this is a contradiction, since we assumed that there are two distinct solutions. Hence the inequality is strict. As a result, we can write equation (B.38) as implying:

$$M_1 < D_{i,1} \hat{W}^{\rho_1} M_1^{1^{+}\left(\frac{a}{a+1}+b_{11}\right)} m_1^{1^{-}\left(\frac{a}{a+1}+b_{11}\right)} M_2^{1^{+}\left(-\frac{1}{a+1}+b_{12}\right)} m_2^{1^{+}\left(-\frac{1}{a+1}+b_{12}\right)} + \sum_j F_{ij,1} M_1^{\frac{1}{a+1}} m_2^{-\frac{a}{a+1}}, \tag{B.39}$$

where the (apologetically cumbersome) notation $1^{+}(x) \equiv \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$ and $1^{-}(x) \equiv \begin{cases} x & \text{if } x < 0 \\ 0 & \text{if } x \geq 0 \end{cases}$, is necessary to consider multiple cases of the signs of the exponents at once.

We can proceed similarly for the minimum bound of equation (B.36), yielding:

$$m_1 > D_{i,1} \hat{W}^{\rho_1} m_1^{1^{+}\left(\frac{a}{a+1}+b_{11}\right)} M_1^{1^{-}\left(\frac{a}{a+1}+b_{11}\right)} m_2^{1^{+}\left(-\frac{1}{a+1}+b_{12}\right)} M_2^{1^{+}\left(-\frac{1}{a+1}+b_{12}\right)} + \sum_j F_{ij,1} m_1^{\frac{1}{a+1}} M_2^{-\frac{a}{a+1}}. \tag{B.40}$$

Taking ratios of equations (B.39) and (B.40) yields:

$$\begin{aligned}
&\frac{M_1}{m_1} < \frac{D_{i,1}\hat{W}^{\rho_1} M_1^{1^+\left(\frac{a}{a+1}+b_{11}\right)} m_1^{1^-\left(\frac{a}{a+1}+b_{11}\right)} M_2^{1^+\left(-\frac{1}{a+1}+b_{12}\right)} m_2^{1^+\left(-\frac{1}{a+1}+b_{12}\right)} + \sum_j F_{ij,1} M_1^{\frac{1}{a+1}} m_2^{-\frac{a}{a+1}}}{D_{i,1}\hat{W}^{\rho_1} m_1^{1^+\left(\frac{a}{a+1}+b_{11}\right)} M_1^{1^-\left(\frac{a}{a+1}+b_{11}\right)} m_2^{1^+\left(-\frac{1}{a+1}+b_{12}\right)} M_2^{1^+\left(-\frac{1}{a+1}+b_{12}\right)} + \sum_j F_{ij,1} m_1^{\frac{1}{a+1}} M_2^{-\frac{a}{a+1}}} \iff \\
&\mu_1 < \beta_1 \mu_1^{\left|\frac{a}{a+1}+b_{11}\right|} \mu_2^{\left|-\frac{1}{a+1}+b_{12}\right|} + (1-\beta_1)\, \mu_1^{\frac{1}{a+1}} \mu_2^{\frac{a}{a+1}},
\end{aligned} \tag{B.41}$$

where $\beta_1 \equiv \frac{D_{i,1}\hat{W}^{\rho_1} m_1^{1^+\left(\frac{a}{a+1}+b_{11}\right)} M_1^{1^-\left(\frac{a}{a+1}+b_{11}\right)} m_2^{1^+\left(-\frac{1}{a+1}+b_{12}\right)} M_2^{1^+\left(-\frac{1}{a+1}+b_{12}\right)}}{D_{i,1}\hat{W}^{\rho_1} m_1^{1^+\left(\frac{a}{a+1}+b_{11}\right)} M_1^{1^-\left(\frac{a}{a+1}+b_{11}\right)} m_2^{1^+\left(-\frac{1}{a+1}+b_{12}\right)} M_2^{1^+\left(-\frac{1}{a+1}+b_{12}\right)} + \sum_j F_{ij,1} m_1^{\frac{1}{a+1}} M_2^{-\frac{a}{a+1}}}$.

Proceeding identically for equation (B.37) yields the corresponding bound:

$$\mu_2 < \beta_2 \mu_1^{\left|-\frac{1}{a+1}+b_{21}\right|} \mu_2^{\left|\frac{a}{a+1}+b_{22}\right|} + (1-\beta_2)\, \mu_1^{\frac{a}{a+1}} \mu_2^{\frac{1}{a+1}}, \tag{B.42}$$

where $\beta_2 \equiv \frac{D_{i,2}\hat{W}^{\rho_2} m_1^{1^+\left(-\frac{1}{a+1}+b_{21}\right)} M_1^{1^-\left(-\frac{1}{a+1}+b_{21}\right)} m_2^{1^+\left(\frac{a}{a+1}+b_{22}\right)} M_2^{1^+\left(\frac{a}{a+1}+b_{22}\right)}}{D_{i,2}\hat{W}^{\rho_2} m_1^{1^+\left(-\frac{1}{a+1}+b_{21}\right)} M_1^{1^-\left(-\frac{1}{a+1}+b_{21}\right)} m_2^{1^+\left(\frac{a}{a+1}+b_{22}\right)} M_2^{1^+\left(\frac{a}{a+1}+b_{22}\right)} + \sum_j F_{ij,2} M_1^{-\frac{a}{a+1}} m_2^{\frac{1}{a+1}}}$.

Both equations (B.41) and (B.42) bound μ_l above by a weighted average of two terms. For each equation to be true, it must then be the case that either μ_l is bounded above by the first term or it is bounded above by the second term (or both). Considering all possible combinations, for there not to be a contradiction, we require at least one of the following conditions to be true:

$$\begin{pmatrix}\ln\mu_1\\ \ln\mu_2\end{pmatrix} < \begin{pmatrix}\left|\frac{a}{a+1}+b_{11}\right| & \left|-\frac{1}{a+1}+b_{12}\right| \\ \left|-\frac{1}{a+1}+b_{21}\right| & \left|\frac{a}{a+1}+b_{22}\right|\end{pmatrix}\begin{pmatrix}\ln\mu_1\\ \ln\mu_2\end{pmatrix}$$

$$\begin{pmatrix}\ln\mu_1\\ \ln\mu_2\end{pmatrix} < \begin{pmatrix}\frac{1}{a+1} & \frac{a}{a+1} \\ \frac{a}{a+1} & \frac{1}{a+1}\end{pmatrix}\begin{pmatrix}\ln\mu_1\\ \ln\mu_2\end{pmatrix}$$

$$\begin{pmatrix}\ln\mu_1\\ \ln\mu_2\end{pmatrix} < \begin{pmatrix}\left|\frac{a}{a+1}+b_{11}\right| & \left|-\frac{1}{a+1}+b_{12}\right| \\ \frac{a}{a+1} & \frac{1}{a+1}\end{pmatrix}\begin{pmatrix}\ln\mu_1\\ \ln\mu_2\end{pmatrix}$$

$$\begin{pmatrix}\ln\mu_1\\ \ln\mu_2\end{pmatrix} < \begin{pmatrix}\frac{1}{a+1} & \frac{a}{a+1} \\ \left|-\frac{1}{a+1}+b_{21}\right| & \left|\frac{a}{a+1}+b_{22}\right|\end{pmatrix}\begin{pmatrix}\ln\mu_1\\ \ln\mu_2\end{pmatrix}$$

By the Collatz-Wielandt formula, note the each inequality can hold only if its matrix has a spectral radius greater than one. By the Gershgorin circle theorem, if all row sums of a matrix is no greater than one, its spectral radius is also no greater than one. As a result, none of the four inequalities will hold if:

$$\left|\frac{a}{a+1}+b_{11}\right| + \left|-\frac{1}{a+1}+b_{12}\right| \leq 1 \tag{B.43}$$

and:

$$\left|-\frac{1}{a+1}+b_{21}\right| + \left|\frac{a}{a+1}+b_{22}\right| \leq 1 \tag{B.44}$$

Hence if both inequalities (B.43) and (B.44) are satisfied, then we have a contradiction, thereby establishing uniqueness. Recall that $a = \frac{\theta\lambda}{1+\theta\lambda}$, $b_{11} \equiv \frac{a}{a^2-1}\frac{1}{1-\theta\alpha}$, $b_{12} \equiv -\frac{1}{a^2-1}\frac{1}{1-\theta\alpha}$, $b_{21} \equiv -\frac{1}{a^2-1}\frac{1}{1-\theta\beta}$, and $b_{22} \equiv \frac{a}{a^2-1}\frac{1}{1-\theta\beta}$. Although the calculations are tedious, for all $\alpha \in [0,1]$, $\beta \in [0,1]$, $\theta \geq 0$,and $\lambda \geq 0$, one can verify that inequality (B.43) holds if:

$$\alpha \leq \frac{1}{2}\left(\frac{1}{\theta} - \lambda\right).$$

Similarly, one can verify that inequality (B.44) holds if:

$$\beta \leq \frac{1}{2}\left(\frac{1}{\theta} - \lambda\right),$$

as required.

C Online Appendix: Additional Derivations

In this subsection, we present detailed derivations from the main text.

C.1 Section 2.1: An Economic Geography Model with Optimal Routing

Below, we derive the equilibrium conditions for the economic geography model described in equations (10) and (11) in the paper. We start with the first market clearing condition defined in equation (4) and combine it with the gravity equation described in equation (5):

$$
\begin{aligned}
Y_i &= \sum_{j=1}^{N} X_{ij} \iff \\
Y_i &= \sum_{j=1}^{N} \tau_{ij}^{-\theta} w_i^{-\theta} A_i^{\theta} E_j P_j^{\theta} \iff \\
\frac{Y_i}{A_i^{\theta}} w_i^{\theta} &= \sum_{j=1}^{N} \tau_{ij}^{-\theta} E_j P_j^{\theta} \iff \\
\bar{A}_i^{-\theta} L_i^{1-\alpha\theta} w_i^{\theta+1} &= \sum_{j=1}^{N} \tau_{ij}^{-\theta} w_j L_j P_j^{\theta}.
\end{aligned}
$$

Assuming welfare equalization, the above becomes:

$$
\begin{aligned}
\bar{A}_i^{-\theta} L_i^{1-\alpha\theta} w_i^{\theta+1} &= \sum_{j=1}^{N} \tau_{ij}^{-\theta} w_j L_j w_j^{\theta} \bar{u}_j^{\theta} L_j^{\beta\theta} W^{-\theta} \iff \\
\bar{A}_i^{-\theta} L_i^{1-\alpha\theta} w_i^{\theta+1} &= \sum_{j=1}^{N} \tau_{ij}^{-\theta} w_j^{\theta+1} \bar{u}_j^{\theta} L_j^{\beta\theta+1} W^{-\theta}.
\end{aligned}
$$

Now, defining $y_i = \frac{Y_i}{Y^W} = \frac{w_i L_i}{Y^W}$ and $l_i = \frac{L_i}{\bar{L}}$

$$
\begin{aligned}
\bar{A}_i^{-\theta} L_i^{1-\alpha\theta} w_i^{\theta+1} &= \sum_{j=1}^{N} \tau_{ij}^{-\theta} w_j^{\theta+1} \bar{u}_j^{\theta} L_j^{\beta\theta+1} W^{-\theta} \iff \\
\bar{A}_i^{-\theta} l_i^{1-\alpha\theta} \bar{L}^{1-\alpha\theta} \left(\frac{y_i Y^W}{l_i \bar{L}}\right)^{\theta+1} &= \sum_{j=1}^{N} \tau_{ij}^{-\theta} \left(\frac{y_i Y^W}{l_i \bar{L}}\right)^{\theta+1} \bar{u}_j^{\theta} L_j^{\beta\theta+1} W^{-\theta} \iff \\
\bar{A}_i^{-\theta} y_i^{\theta+1} l_i^{-\theta(1+\alpha)} \bar{L}^{\theta(1+\alpha)} \left(Y^W\right)^{\theta+1} &= \left(Y^W\right)^{\theta+1} \bar{L}^{\theta(\beta-1)} W^{-\theta} \sum_{j=1}^{N} \tau_{ij}^{-\theta} \bar{u}_j^{\theta} y_j^{1+\theta} l_j^{\theta(\beta-1)} \iff \\
\bar{A}_i^{-\theta} y_i^{1+\theta} l_i^{-\theta(1+\alpha)} &= \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}} \sum_{j=1}^{N} \tau_{ij}^{-\theta} \bar{u}_j^{\theta} {y_j}^{1+\theta} {l_j}^{\theta(\beta-1)}.
\end{aligned}
$$

This concludes our derivation of the equilibrium condition we describe in equation (10).

Moving on to the second market clearing condition defined in equation (4) and combining with the gravity equation described in equation (5):

$$
\begin{aligned}
E_i &= \sum_{j=1}^{N} X_{ji} \iff \\
E_i &= \sum_{j=1}^{N} \tau_{ji}^{-\theta} w_j^{-\theta} A_j^{\theta} E_i P_i^{\theta} \iff \\
P_i^{-\theta} &= \sum_{j=1}^{N} \tau_{ji}^{-\theta} w_j^{-\theta} \bar{A}_j^{\theta} L_j^{\alpha\theta}.
\end{aligned}
$$

Assuming welfare equalization, this definition of the consumer price index becomes:

$$
W^{\theta} w_i^{-\theta} \bar{u}_i^{-\theta} L_i^{-\beta\theta} = \sum_{j=1}^{N} \tau_{ji}^{-\theta} w_j^{-\theta} \bar{A}_j^{\theta} L_j^{\alpha\theta}.
$$

Defining the same income and labor shares, $y_i = \frac{Y_i}{Y^W} = \frac{w_i L_i}{Y^W}$ and $l_i = \frac{L_i}{\bar{L}}$, we used for the first equilibrium condition, we get the following:

$$
\begin{aligned}
W^{\theta} w_i^{-\theta} \bar{u}_i^{-\theta} L_i^{-\beta\theta} &= \sum_{j=1}^{N} \tau_{ji}^{-\theta} w_j^{-\theta} \bar{A}_j^{\theta} L_j^{\alpha\theta} \iff \\
W^{\theta} \left(\frac{y_i Y^W}{l_i \bar{L}} \right)^{-\theta} \bar{u}_i^{-\theta} l_i^{-\beta\theta} \bar{L}^{-\beta\theta} &= \sum_{j=1}^{N} \tau_{ji}^{-\theta} \left(\frac{y_j Y^W}{l_j \bar{L}} \right)^{-\theta} \bar{A}_j^{\theta} l_j^{\alpha\theta} \bar{L}^{\alpha\theta} \iff \\
W^{\theta} y_i^{-\theta} l_i^{\theta(1-\beta)} \bar{u}_i^{-\theta} \left(Y^W\right)^{-\theta} \bar{L}^{-\beta\theta} &= \left(Y^W\right)^{-\theta} \bar{L}^{\theta(\alpha+1)} \sum_{j=1}^{N} \tau_{ji}^{-\theta} y_j^{-\theta} l_j^{\theta(\alpha+1)} \bar{A}_j^{\theta} \iff \\
\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)} &= \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}} \sum_{j=1}^{N} \tau_{ji}^{-\theta} \bar{A}_j^{\theta} y_j^{-\theta} l_j^{\theta(\alpha+1)}.
\end{aligned}
$$

At this point, we have reached the second equilibrium condition described in equation (11).

C.2 Section 2.2: An Urban Model with Optimal Routing

In this section, we derive the equilibrium conditions for the urban model described in equations (19) and (20). We start by combining commuting gravity with the adding-up constraint on the residential population:

$$
\begin{aligned}
L_i^R &= \sum_j L_{ij} \iff \\
&= \sum_j \tau_{ij}^{-\theta} u_i^{\theta} A_j^{\theta} \left(\frac{\bar{L}}{W^{\theta}} \right).
\end{aligned}
$$

We substitute in for the spillovers as defined in equation (18):

$$
L_i^R = \sum_j \tau_{ij}^{-\theta} \bar{u}_i^{\theta} \left(L_i^R\right)^{\beta\theta} \bar{A}_j^{\theta} \left(L_j^F\right)^{\alpha\theta} \left(\frac{\bar{L}}{W^{\theta}} \right).
$$

Next, we define residential labor shares and commercial labor shares as $l_i^R = L_i^R/\bar{L}, l_i^F = L_i^F/\bar{L}$, and putting the above equation in terms of the shares, we get equilibrium equation (19):

$$
\begin{aligned}
L_i^R &= \sum_j \tau_{ij}^{-\theta} \bar{u}_i^{\theta} \left(L_i^R\right)^{\beta\theta} \bar{A}_j^{\theta} \left(L_j^F\right)^{\alpha\theta} \left(\frac{\bar{L}}{W^{\theta}}\right) \iff \\
l_i^R \bar{L} &= \sum_j \tau_{ij}^{-\theta} \bar{u}_i^{\theta} \left(l_i^R \bar{L}\right)^{\beta\theta} \bar{A}_j^{\theta} \left(l_j^F \bar{L}\right)^{\alpha\theta} \left(\frac{\bar{L}}{W^{\theta}}\right) \iff \\
\left(l_i^R\right)^{1-\beta\theta} &= \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}} \sum_j \tau_{ij}^{-\theta} \bar{u}_i^{\theta} \bar{A}_j^{\theta} \left(l_j^F\right)^{\alpha\theta}.
\end{aligned}
$$

Moving to the derivation of the second equilibrium condition, we start by combining commuting gravity with the adding up constraint on commercial population:

$$
\begin{aligned}
L_i^F &= \sum_j L_{ji} \iff \\
&= \sum_j \tau_{ji}^{-\theta} u_j^{\theta} A_i^{\theta} \left(\frac{\bar{L}}{W^{\theta}}\right).
\end{aligned}
$$

We substitute in for the spillovers to arrive at the following characterization of the commercial labor force:

$$
L_i^F = \sum_j \tau_{ji}^{-\theta} \bar{u}_j^{\theta} \left(L_j^R\right)^{\beta\theta} \bar{A}_i^{\theta} \left(L_j^F\right)^{\alpha\theta} \left(\frac{\bar{L}}{W^{\theta}}\right).
$$

Finally, we define the above expression in terms of residential and commercial labor shares:

$$
\begin{aligned}
L_i^F &= \sum_j \tau_{ji}^{-\theta} \bar{u}_j^{\theta} \left(L_j^R\right)^{\beta\theta} \bar{A}_i^{\theta} \left(L_j^F\right)^{\alpha\theta} \left(\frac{\bar{L}}{W^{\theta}}\right) \iff \\
l_i^F \bar{L} &= \sum_j \tau_{ji}^{-\theta} \bar{u}_j^{\theta} \left(l_j^R \bar{L}\right)^{\beta\theta} \bar{A}_i^{\theta} \left(l_j^F \bar{L}\right)^{\alpha\theta} \left(\frac{\bar{L}}{W^{\theta}}\right) \iff \\
\left(l_i^F\right)^{1-\alpha\theta} &= \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}} \sum_j \tau_{ji}^{-\theta} \bar{A}_i^{\theta} \bar{u}_j^{\theta} \left(l_j^R\right)^{\beta\theta},
\end{aligned}
$$

which is the second equilibrium condition defined in equation (20).

C.3 Section 4.1: Equilibrium

Trade Model In this Appendix section, we derive the equilibrium conditions for the economic geography and commuting frameworks.

For the trade equilibrium conditions, we start with equation (10) from the paper. Note that $\tau_{ij}^{-\theta} = [\mathbf{I}-\mathbf{A}]_{ij}^{-1}$, where $\mathbf{A} = [a_{ij}] = \left[t_{ij}^{-\theta}\right]$ is the adjacency matrix, so with a change of notation, we can rewrite the summation term as a matrix product:

$$
\begin{aligned}
\left[\bar{A}_i^{-\theta} y_i^{1+\theta} l_i^{-\theta(1+\alpha)}\right] &= \frac{L^{(\alpha+\beta)\theta}}{W^{\theta}} \times \left[\tau_{ij}^{-\theta}\right] \times \left[\bar{u}_j^{\theta} {y_j}^{1+\theta} {l_j}^{\theta(\beta-1)}\right] \iff \\
\left[\bar{A}_i^{-\theta} y_i^{1+\theta} l_i^{-\theta(1+\alpha)}\right] &= \frac{L^{(\alpha+\beta)\theta}}{W^{\theta}} \times [\mathbf{I}-\mathbf{A}]^{-1} \times \left[\bar{u}_j^{\theta} {y_j}^{1+\theta} {l_j}^{\theta(\beta-1)}\right]
\end{aligned}
$$

where $\left[\bar{A}_i^{-\theta} y_i^{1+\theta} l_i^{-\theta(1+\alpha)}\right]$ and $\left[\bar{u}_j^{\theta} {y_j}^{1+\theta} {l_j}^{\theta(\beta-1)}\right]$ are column vectors. Taking a matrix inversion and converting back to summation notation:

$$\begin{aligned}
[\mathbf{I}-\mathbf{A}] \times \left[\bar{A}_i^{-\theta} y_i^{1+\theta} l_i^{-\theta(1+\alpha)}\right] &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta} \times \left[\bar{u}_i^\theta {y_i}^{1+\theta} {l_i}^{\theta(\beta-1)}\right] \iff \\
\left[\bar{A}_i^{-\theta} y_i^{1+\theta} l_i^{-\theta(1+\alpha)}\right] - \mathbf{A} \times \left[\bar{A}_j^{-\theta} y_j^{1+\theta} l_j^{-\theta(1+\alpha)}\right] &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta} \times \left[\bar{u}_i^\theta {y_i}^{1+\theta} {l_i}^{\theta(\beta-1)}\right] \iff \\
\bar{A}_i^{-\theta} y_i^{1+\theta} l_i^{-\theta(1+\alpha)} - \sum_j a_{ij} \bar{A}_j^{-\theta} y_j^{1+\theta} l_j^{-\theta(1+\alpha)} &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta} \bar{u}_i^\theta {y_i}^{1+\theta} {l_i}^{\theta(\beta-1)}
\end{aligned}$$

The second equilibrium condition, equation (11), can also be written as a matrix multiplication, where $\left[\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)}\right]$ and $\left[\bar{A}_j^{\theta} y_j^{-\theta} l_j^{\theta(\alpha+1)}\right]$ are row vectors. Applying the same matrix inversion we did to equilibrium equation we did to the first equilibrium condition, we get:

$$\begin{aligned}
\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)} &= \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^\theta} \sum_{j=1}^N \tau_{ji}^{-\theta} \bar{A}_j^\theta y_j^{-\theta} l_j^{\theta(\alpha+1)} \iff \\
\left[\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)}\right] &= \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^\theta} \times \left[\bar{A}_j^\theta y_j^{-\theta} l_j^{\theta(\alpha+1)}\right] \times \left[\tau_{ji}^{-\theta}\right] \iff \\
\left[\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)}\right] &= \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^\theta} \times \left[\bar{A}_j^\theta y_j^{-\theta} l_j^{\theta(\alpha+1)}\right] \times \left[\mathbf{I}-\mathbf{A}^T\right]^{-1} \iff \\
\left[\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)}\right] \times \left[\mathbf{I}-\mathbf{A}^T\right] &= \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^\theta} \times \left[\bar{A}_i^\theta y_i^{-\theta} l_i^{\theta(\alpha+1)}\right] \iff \\
\left[\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)}\right] - \left[\bar{u}_j^{-\theta} y_j^{-\theta} l_j^{\theta(1-\beta)}\right] \times \mathbf{A}^T &= \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^\theta} \times \left[\bar{A}_i^\theta y_i^{-\theta} l_i^{\theta(\alpha+1)}\right] \iff \\
\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)} - \sum_j a_{ji} \bar{u}_j^{-\theta} y_j^{-\theta} l_j^{\theta(1-\beta)} &= \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^\theta} \bar{A}_i^\theta y_i^{-\theta} l_i^{\theta(\alpha+1)}
\end{aligned}$$

Recalling that $a_{ij} \equiv t_{ij}^{-\theta}$, we have for our two equilibrium conditions (before incorporating traffic congestion):

$$\begin{aligned}
\bar{A}_i^{-\theta} y_i^{1+\theta} l_i^{-\theta(1+\alpha)} &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta} \bar{u}_i^\theta {y_i}^{1+\theta} {l_i}^{\theta(\beta-1)} + \sum_j t_{ij}^{-\theta} \bar{A}_j^{-\theta} y_j^{1+\theta} l_j^{-\theta(1+\alpha)} \\
\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)} &= \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^\theta} \bar{A}_i^\theta y_i^{-\theta} l_i^{\theta(\alpha+1)} + \sum_j t_{ji}^{-\theta} \bar{u}_j^{-\theta} y_j^{-\theta} l_j^{\theta(1-\beta)}
\end{aligned}$$

To incorporate congestion, we combine these two equations with the expression (26), converting from market access terms to equilibrium $\{y_i\}$ and $\{l_i\}$ (as in Appendix B). Starting with the first equilibrium condition:

$$
\begin{aligned}
\bar{A}_i^{-\theta} y_i^{1+\theta} l_i^{-\theta(1+\alpha)} =& \frac{L^{(\alpha+\beta)\theta}}{W^{\theta}} \bar{u}_i^{\theta} {y_i}^{1+\theta} {l_i}^{\theta(\beta-1)} + \\
& \sum_j \left(\left(\bar{t}_{ij} \bar{L}^{\lambda}\right)^{\frac{1}{1+\theta\lambda}} \left(\frac{\bar{L}^{-(\alpha+\beta)\theta}}{W^{-\theta}} \right)^{\frac{\lambda}{1+\theta\lambda}} \bar{A}_j^{-\frac{\theta\lambda}{1+\theta\lambda}} \bar{u}_i^{-\frac{\theta\lambda}{1+\theta\lambda}} l_i^{-\frac{\theta\lambda(\beta-1)}{1+\theta\lambda}} l_j^{-\frac{\theta\lambda(1+\alpha)}{1+\theta\lambda}} y_i^{-\frac{\theta\lambda}{1+\theta\lambda}} y_j^{\frac{\lambda(1+\theta)}{1+\theta\lambda}} \right)^{-\theta} \times \\
& \bar{A}_j^{-\theta} y_j^{1+\theta} l_j^{-\theta(1+\alpha)} \iff \\
\bar{A}_i^{-\theta} y_i^{1+\theta} l_i^{-\theta(1+\alpha)} =& \frac{L^{(\alpha+\beta)\theta}}{W^{\theta}} \bar{u}_i^{\theta} {y_i}^{1+\theta} {l_i}^{\theta(\beta-1)} + \\
& \sum_j \left(\bar{t}_{ij} \bar{L}^{\lambda}\right)^{-\frac{\theta}{1+\theta\lambda}} \left(\frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}} \right)^{\frac{\theta\lambda}{1+\theta\lambda}} \bar{A}_j^{-\frac{\theta}{1+\theta\lambda}} \bar{u}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} l_i^{\theta(\beta-1)\frac{\theta\lambda}{1+\theta\lambda}} y_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} y_j^{\frac{1+\theta}{1+\theta\lambda}} l_j^{-\frac{\theta(1+\alpha)}{1+\theta\lambda}} \iff \\
y_i^{\frac{1+\theta\lambda+\theta}{1+\theta\lambda}} l_i^{\frac{-\theta(1+\alpha+\theta\lambda(\alpha+\beta))}{1+\theta\lambda}} =& \frac{L^{(\alpha+\beta)\theta}}{W^{\theta}} \bar{A}_i^{\theta} \bar{u}_i^{\theta} y_i^{\frac{1+\theta\lambda+\theta}{1+\theta\lambda}} l_i^{\frac{\theta(\beta-1)}{1+\theta\lambda}} + \\
& \left(\frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}} \right)^{\frac{\theta\lambda}{1+\theta\lambda}} \sum_j \left(\bar{t}_{ij} \bar{L}^{\lambda}\right)^{-\frac{\theta}{1+\theta\lambda}} \times \bar{A}_i^{\theta} \bar{u}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{A}_j^{-\frac{\theta}{1+\theta\lambda}} y_j^{\frac{1+\theta}{1+\theta\lambda}} l_j^{-\frac{\theta(1+\alpha)}{1+\theta\lambda}} \iff \\
y_i^{\frac{1+\theta\lambda+\theta}{1+\theta\lambda}} l_i^{\frac{-\theta(1+\alpha+\theta\lambda(\alpha+\beta))}{1+\theta\lambda}} =& \chi \bar{A}_i^{\theta} \bar{u}_i^{\theta} y_i^{\frac{1+\theta\lambda+\theta}{1+\theta\lambda}} l_i^{\frac{\theta(\beta-1)}{1+\theta\lambda}} + \chi^{\frac{\theta\lambda}{1+\theta\lambda}} \sum_j \left(\bar{t}_{ij} \bar{L}^{\lambda}\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{A}_i^{\theta} \bar{u}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{A}_j^{-\frac{\theta}{1+\theta\lambda}} y_j^{\frac{1+\theta}{1+\theta\lambda}} l_j^{-\frac{\theta(1+\alpha)}{1+\theta\lambda}},
\end{aligned}
$$

where $\chi = \left(\frac{L^{\alpha+\beta}}{W}\right)^{\theta}$, as in equation (28). For the second equilibrium condition, we proceed similarly:

$$
\begin{aligned}
\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)} =& \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}} \bar{A}_i^{\theta} y_i^{-\theta} l_i^{\theta(\alpha+1)} + \\
& \sum_j \left(\left(\bar{t}_{ji} \bar{L}^{\lambda}\right)^{\frac{1}{1+\theta\lambda}} \left(\frac{\bar{L}^{-(\alpha+\beta)\theta}}{W^{-\theta}} \right)^{\frac{\lambda}{1+\theta\lambda}} \bar{A}_i^{-\frac{\theta\lambda}{1+\theta\lambda}} \bar{u}_j^{-\frac{\theta\lambda}{1+\theta\lambda}} l_j^{-\frac{\theta\lambda(\beta-1)}{1+\theta\lambda}} l_i^{-\frac{\theta\lambda(1+\alpha)}{1+\theta\lambda}} y_j^{-\frac{\theta\lambda}{1+\theta\lambda}} y_i^{\frac{\lambda(1+\theta)}{1+\theta\lambda}} \right)^{-\theta} \times \\
& \bar{u}_j^{-\theta} y_j^{-\theta} l_j^{\theta(1-\beta)} \iff \\
\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)} =& \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}} \bar{A}_i^{\theta} y_i^{-\theta} l_i^{\theta(\alpha+1)} + \\
& \sum_j \left(\bar{t}_{ji} \bar{L}^{\lambda}\right)^{-\frac{\theta}{1+\theta\lambda}} \times \left(\frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}} \right)^{\frac{\theta\lambda}{1+\theta\lambda}} \bar{A}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{u}_j^{-\frac{\theta}{1+\theta\lambda}} l_j^{\frac{\theta(1-\beta)}{1+\theta\lambda}} l_i^{\theta(1+\alpha)\frac{\theta\lambda}{1+\theta\lambda}} y_j^{-\frac{\theta}{1+\theta\lambda}} y_i^{-\frac{\theta\lambda(1+\theta)}{1+\theta\lambda}} \iff \\
y_i^{-\frac{\theta(1-\lambda)}{1+\theta\lambda}} l_i^{\frac{\theta(1-\beta-\theta\lambda(\alpha+\beta))}{1+\theta\lambda}} =& \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}} \bar{A}_i^{\theta} \bar{u}_i^{\theta} y_i^{-\frac{\theta(1-\lambda)}{1+\theta\lambda}} l_i^{\frac{\theta(\alpha+1)}{1+\theta\lambda}} + \\
& \left(\frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}} \right)^{\frac{\theta\lambda}{1+\theta\lambda}} \sum_j \left(\bar{t}_{ji} \bar{L}^{\lambda}\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{A}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{u}_i^{\theta} \bar{u}_j^{-\frac{\theta}{1+\theta\lambda}} y_j^{-\frac{\theta}{1+\theta\lambda}} l_j^{\frac{\theta(1-\beta)}{1+\theta\lambda}} \iff \\
y_i^{-\frac{\theta(1-\lambda)}{1+\theta\lambda}} l_i^{\frac{\theta(1-\beta-\theta\lambda(\alpha+\beta))}{1+\theta\lambda}} =& \chi \bar{A}_i^{\theta} \bar{u}_i^{\theta} y_i^{-\frac{\theta(1-\lambda)}{1+\theta\lambda}} l_i^{\frac{\theta(\alpha+1)}{1+\theta\lambda}} + \chi^{\frac{\theta\lambda}{1+\theta\lambda}} \sum_j \left(\bar{t}_{ji} \bar{L}^{\lambda}\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{A}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{u}_i^{\theta} \bar{u}_j^{-\frac{\theta}{1+\theta\lambda}} y_j^{-\frac{\theta}{1+\theta\lambda}} l_j^{\frac{\theta(1-\beta)}{1+\theta\lambda}}
\end{aligned}
$$

where again $\chi = \left(\frac{L^{\alpha+\beta}}{W}\right)^{\theta}$, as in equation (29).

Commuting Model The derivations for the commuting model follow a very similar process to that of the economic geography model. We rewrite the first equilibrium condition, equation (19), as a matrix product and invert:

$$\begin{aligned}
\left[\bar{u}_i^{-\theta}\left(l_i^R\right)^{-\theta\beta+1}\right] &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\times\left[\tau_{ij}^{-\theta}\right]\times\left[\bar{A}_j^\theta\left(l_j^F\right)^{\theta\alpha}\right] \iff \\
\left[\bar{u}_i^{-\theta}\left(l_i^R\right)^{-\theta\beta+1}\right] &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\times\left[\mathbf{I}-\mathbf{A}\right]^{-1}\times\left[\bar{A}_j^\theta\left(l_j^F\right)^{\theta\alpha}\right] \\
\left[\mathbf{I}-\mathbf{A}\right]\times\left[\bar{u}_i^{-\theta}\left(l_i^R\right)^{-\theta\beta+1}\right] &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\times\left[\bar{A}_j^\theta\left(l_j^F\right)^{\theta\alpha}\right] \iff \\
\left[\bar{u}_i^{-\theta}\left(l_i^R\right)^{-\theta\beta+1}\right]-\mathbf{A}\times\left[\bar{u}_j^{-\theta}\left(l_j^R\right)^{-\theta\beta+1}\right] &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\times\left[\bar{A}_i^\theta\left(l_i^F\right)^{\theta\alpha}\right] \iff \\
\bar{u}_i^{-\theta}\left(l_i^R\right)^{-\theta\beta+1}-\sum_j a_{ij}\bar{u}_j^{-\theta}\left(l_j^R\right)^{-\theta\beta+1} &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\bar{A}_i^\theta\left(l_i^F\right)^{\theta\alpha}
\end{aligned}$$

where $\left[\bar{u}_i^\theta\left(l_i^R\right)^{-\theta\beta+1}\right]$ and $\left[T_j^\theta\left(l_j^F\right)^{\theta\alpha}\right]$ are column vectors.

Applying the same steps to equilibrium equation (20), where $\left[\bar{A}_i^{-\theta}\left(l_i^F\right)^{-\theta\alpha+1}\right]$ and $\left[\bar{u}_j^\theta\left(l_j^R\right)^{\theta\beta}\right]$ are row vectors:

$$\begin{aligned}
\left(l_i^F\right)^{-\theta\alpha+1} &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\sum_j \bar{u}_j^\theta\tau_{ji}^{-\theta}\bar{A}_i^\theta\left(l_j^R\right)^{\theta\beta} \iff \\
\left[\bar{A}_i^{-\theta}\left(l_i^F\right)^{-\theta\alpha+1}\right] &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\times\left[\bar{u}_j^\theta\left(l_j^R\right)^{\theta\beta}\right]\times\left[\tau_{ji}^{-\theta}\right] \iff \\
\left[\bar{A}_i^{-\theta}\left(l_i^F\right)^{-\theta\alpha+1}\right] &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\times\left[\bar{u}_j^\theta\left(l_j^R\right)^{\theta\beta}\right]\times\left[\mathbf{I}-\mathbf{A}^T\right]^{-1} \iff \\
\left[\bar{A}_i^{-\theta}\left(l_i^F\right)^{-\theta\alpha+1}\right]\times\left[\mathbf{I}-\mathbf{A}^T\right] &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\times\left[\bar{u}_i^\theta\left(l_i^R\right)^{\theta\beta}\right] \iff \\
\left[\bar{A}_i^{-\theta}\left(l_i^F\right)^{-\theta\alpha+1}\right]-\left[\bar{A}_j^{-\theta}\left(l_j^F\right)^{-\theta\alpha+1}\right]\times\mathbf{A}^T &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\times\left[\bar{u}_i^\theta\left(l_i^R\right)^{\theta\beta}\right] \iff \\
T_i^{-\theta}\left(l_i^F\right)^{-\theta\alpha+1}-\sum_j a_{ji}\bar{A}_j^{-\theta}\left(l_j^F\right)^{-\theta\alpha+1} &= \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\bar{u}_i^\theta\left(l_i^R\right)^{\theta\beta}
\end{aligned}$$

Recalling $a_{ij}\equiv t_{ij}^{-\theta}$,we have two equilibrium conditions for our commuting model:

$$\bar{u}_i^{-\theta}\left(l_i^R\right)^{-\theta\beta+1} = \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\bar{A}_i^\theta\left(l_i^F\right)^{\theta\alpha}+\sum_j t_{ij}^{-\theta}\bar{u}_j^{-\theta}\left(l_j^R\right)^{-\theta\beta+1}$$

$$\bar{A}_i^{-\theta}\left(l_i^F\right)^{-\theta\alpha+1} = \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\bar{u}_i^\theta\left(l_i^R\right)^{\theta\beta}+\sum_j t_{ji}^{-\theta}\bar{A}_j^{-\theta}\left(l_j^F\right)^{-\theta\alpha+1}$$

As above, substituting in our expression for the iceberg transportation costs along a link using equation (26), and converting from market access terms to equilibrium $\left\{l_i^F\right\}$ and $\left\{l_i^R\right\}$ (as in Appendix B), incorporates endogenous traffic congestion. For the first equilibrium condition (30), we have:

$$
\begin{aligned}
\bar{u}_i^{-\theta}\left(l_i^R\right)^{-\theta\beta+1} =& \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\bar{A}_i^\theta\left(l_i^F\right)^{\theta\alpha} + \\
& \sum_j \left(\left(\bar{t}_{ij}\bar{L}^\lambda\right)^{\frac{1}{1+\theta\lambda}} \bar{A}_i^{\frac{-\theta\lambda}{1+\theta\lambda}} \left(l_i^F\right)^{\frac{(1-\alpha\theta)\lambda}{1+\theta\lambda}} \bar{u}_j^{\frac{-\theta\lambda}{1+\theta\lambda}} \left(l_j^R\right)^{\frac{(1-\beta\theta)\lambda}{1+\theta\lambda}} \times W^{\frac{\theta\lambda}{1+\theta\lambda}} \bar{L}^{\frac{-\theta\lambda(\alpha+\beta)}{1+\theta\lambda}}\right)^{-\theta} \bar{u}_j^{-\theta}\left(l_j^R\right)^{-\theta\beta+1} \iff \\
\bar{u}_i^{-\theta}\left(l_i^R\right)^{-\theta\beta+1} =& \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\bar{A}_i^\theta\left(l_i^F\right)^{\theta\alpha} + \\
& \sum_j \left(\bar{t}_{ij}\bar{L}^\lambda\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{u}_j^{\theta\frac{\theta\lambda}{1+\theta\lambda}-\theta} \left(l_j^R\right)^{-\frac{\theta\lambda(1-\beta\theta)}{1+\theta\lambda}+(1-\beta\theta)} \bar{A}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} \left(l_i^F\right)^{-\frac{\theta\lambda(1-\alpha\theta)}{1+\theta\lambda}} W^{-\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{L}^{\frac{\theta\lambda(\alpha\theta+\beta\theta)}{1+\theta\lambda}} \iff \\
\left(l_i^R\right)^{-\theta\beta+1}\left(l_i^F\right)^{\frac{\theta\lambda(1-\alpha\theta)}{1+\theta\lambda}} =& \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\bar{A}_i^\theta\bar{u}_i^\theta\left(l_i^F\right)^{\theta\alpha+\frac{\theta\lambda(1-\alpha\theta)}{1+\theta\lambda}} + \\
& \sum_j \left(\bar{t}_{ij}\bar{L}^\lambda\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{u}_j^{\theta\frac{\theta\lambda}{1+\theta\lambda}-\theta} \bar{u}_i^\theta \left(l_j^R\right)^{-\frac{\theta\lambda(1-\beta\theta)}{1+\theta\lambda}+(1-\beta\theta)} \bar{A}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} W^{-\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{L}^{\frac{\theta\lambda(\alpha\theta+\beta\theta)}{1+\theta\lambda}} \iff \\
\left(l_i^R\right)^{-\theta\beta+1}\left(l_i^F\right)^{\frac{\theta\lambda(1-\alpha\theta)}{1+\theta\lambda}} =& \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\bar{A}_i^\theta\bar{u}_i^\theta\left(l_i^F\right)^{\theta\alpha+\frac{\theta\lambda(1-\alpha\theta)}{1+\theta\lambda}} + \\
& \sum_j \left(\bar{t}_{ij}\bar{L}^\lambda\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{u}_j^{\theta\frac{\theta\lambda}{1+\theta\lambda}-\theta} \bar{u}_i^\theta \left(l_j^R\right)^{-\frac{\theta\lambda(1-\beta\theta)}{1+\theta\lambda}+(1-\beta\theta)} \bar{A}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} W^{-\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{L}^{\frac{\theta\lambda(\alpha\theta+\beta\theta)}{1+\theta\lambda}} \iff \\
\left(l_i^R\right)^{-\theta\beta+1}\left(l_i^F\right)^{\frac{\theta\lambda(1-\alpha\theta)}{1+\theta\lambda}} =& \chi\bar{A}_i^\theta\bar{u}_i^\theta\left(l_i^F\right)^{\frac{\theta(\alpha+\lambda)}{1+\theta\lambda}} + \chi^{\frac{\theta\lambda}{1+\theta\lambda}} \sum_j \left(\bar{t}_{ij}\bar{L}^\lambda\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{A}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{u}_j^{-\frac{\theta}{1+\theta\lambda}} \bar{u}_i^\theta \left(l_j^R\right)^{\frac{1-\beta\theta}{1+\theta\lambda}}
\end{aligned}
$$

where $\chi = \left(\frac{L^{\alpha+\beta}}{W}\right)^\theta$, as claimed. For the second equilibrium condition (31):

$$
\begin{aligned}
\bar{A}_i^{-\theta}\left(l_i^F\right)^{-\theta\alpha+1} =& \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\bar{u}_i^\theta\left(l_i^R\right)^{\theta\beta} + \\
& \sum_j \left(\left(\bar{t}_{ji}\bar{L}^\lambda\right)^{\frac{1}{1+\theta\lambda}} \bar{u}_i^{-\frac{\theta\lambda}{1+\theta\lambda}} \left(l_i^R\right)^{\frac{(1-\beta\theta)\lambda}{1+\theta\lambda}} \bar{A}_j^{-\frac{\theta\lambda}{1+\theta\lambda}} \left(l_j^F\right)^{\frac{(1-\alpha\theta)\lambda}{1+\theta\lambda}} W^{\frac{\theta\lambda}{1+\theta\lambda}} \bar{L}^{-\frac{\theta\lambda(\alpha+\beta)}{1+\theta\lambda}}\right)^{-\theta} \bar{A}_j^{-\theta}\left(l_j^F\right)^{-\theta\alpha+1} \iff \\
\bar{A}_i^{-\theta}\left(l_i^F\right)^{-\theta\alpha+1} =& \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\bar{u}_i^\theta\left(l_i^R\right)^{\theta\beta} + \\
& \sum_j \left(\bar{t}_{ji}\bar{L}^\lambda\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{u}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} \left(l_i^R\right)^{-\frac{\theta\lambda(1-\beta\theta)}{1+\theta\lambda}} \bar{A}_j^{\theta\frac{\theta\lambda}{1+\theta\lambda}-\theta} \left(l_j^F\right)^{-\frac{\theta\lambda(1-\alpha\theta)}{1+\theta\lambda}+(1-\alpha\theta)} W^{-\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{L}^{\frac{\theta\lambda(\alpha\theta+\beta\theta)}{1+\theta\lambda}} \iff \\
\left(l_i^F\right)^{-\theta\alpha+1}\left(l_i^R\right)^{\frac{\theta\lambda(1-\beta\theta)}{1+\theta\lambda}} =& \frac{L^{(\alpha+\beta)\theta}}{W^\theta}\bar{A}_i^\theta\bar{u}_i^\theta\left(l_i^R\right)^{\theta\beta+\frac{\theta\lambda(1-\beta\theta)}{1+\theta\lambda}} + \\
& \sum_j \left(\bar{t}_{ji}\bar{L}^\lambda\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{u}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{A}_i^\theta \bar{A}_j^{\theta\frac{\theta\lambda}{1+\theta\lambda}-\theta} \left(l_j^F\right)^{-\frac{\theta\lambda(1-\alpha\theta)}{1+\theta\lambda}+(1-\alpha\theta)} W^{-\theta\frac{\theta\lambda}{1+\theta\lambda}} \bar{L}^{\frac{\theta\lambda(\alpha\theta+\beta\theta)}{1+\theta\lambda}} \iff \\
\left(l_i^F\right)^{-\theta\alpha+1}\left(l_i^R\right)^{\frac{\theta\lambda(1-\beta\theta)}{1+\theta\lambda}} =& \chi\bar{A}_i^\theta\bar{u}_i^\theta\left(l_i^R\right)^{\frac{\theta(\beta+\lambda)}{1+\theta\lambda}} + \chi^{\frac{\theta\lambda}{1+\theta\lambda}} \sum_j \left(\bar{t}_{ji}\bar{L}^\lambda\right)^{-\frac{\theta}{1+\theta\lambda}} \bar{A}_i^\theta \bar{A}_j^{-\frac{\theta}{1+\theta\lambda}} \bar{u}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}} \left(l_j^F\right)^{\frac{1-\alpha\theta}{1+\theta\lambda}},
\end{aligned}
$$

as claimed.

C.4 Section 4.3: Scale Dependence

In this section, we present a derivation of the partial derivative of trade costs about $c = 1$. Let's define the matrix of bilateral trade costs as $\mathbf{B} \equiv \left(\mathbf{I} - \left(\exp\left(-\theta\ln c\right)\mathbf{A}\right)\right)^{-1}$. We then have from, by matrix calculus, that:

$$\frac{\partial \mathbf{B}}{\partial \ln c} = -\left(\mathbf{B}\right)\frac{\partial\left(\mathbf{I}-\left(\exp\left(-\theta \ln c\right)\mathbf{A}\right)\right)}{\partial \ln c}\left(\mathbf{B}\right) \iff$$
$$\frac{\partial \mathbf{B}}{\partial \ln c} = -\theta\left(\mathbf{B}\right)c^{-\theta}\mathbf{A}\left(\mathbf{B}\right)$$

so that:

$$\left[\frac{\partial \mathbf{B}}{\partial \ln c}\right]_{ij} = -\theta\sum_k\sum_l \mathbf{B}_{ik}\bar{t}_{kl}^{-\theta}\mathbf{B}_{lj}$$

or in our notation:

$$\frac{\partial \tau_{ij}^{-\theta}\left(c\right)}{\partial \ln c}|_{c=1} = -\theta\sum_{k=1}^N\sum_{l=1}^N \tau_{ik}^{-\theta}\bar{t}_{kl}^{-\theta}\tau_{lj}^{-\theta} \iff$$
$$\frac{\partial \ln \tau_{ij}\left(c\right)}{\partial \ln c}|_{c=1} = \sum_{k=1}^N\sum_{l=1}^N \frac{\tau_{ik}^{-\theta}\bar{t}_{kl}^{-\theta}\tau_{lj}^{-\theta}}{\tau_{ij}^{-\theta}}$$

Recall:

$$\pi_{ij}^{kl} = \left(\frac{\tau_{ij}}{\tau_{ik}t_{kl}\tau_{lj}}\right)^{\theta}$$

so that:

$$\frac{\partial \ln \tau_{ij}\left(c\right)}{\partial \ln c}|_{c=1} = \sum_{k=1}^N\sum_{l=1}^N \pi_{ij}^{kl}$$

C.5 Section 5.2: Counterfactuals

To simplify the derivation of counterfactual expressions for the trade and commuting models, we present a generalized version of models in term of market access terms and convert that to the "exact hat" notation developed in Dekle, Eaton, and Kortum (2008). Substituting in for the specific definitions of market access of each model returns the relevant counterfactual equilibrium conditions for both models.

For both models, flows X_{ij} can be described in gravity form as:

$$X_{ij} = \tau_{ij}^{-\theta} \times \frac{\gamma_i}{\Pi_i^{-\theta}} \times \frac{\delta_j}{P_j^{-\theta}}$$

where γ_i and δ_j are cumulative flows out of an origin and into a destination, respectively, and Π_i and P_j are origin and destination market access terms. Trade costs can be represented as:

$$\left[\tau_{ij}^{-\theta}\right] = \left(I - \left[t_{ij}\right]^{-\theta}\right)^{-1}$$

And the iceberg cost of traveling along a link can be described as:

$$t_{ij} = \bar{t}_{ij}\left[\Xi_{ij}\right]^{-\lambda} \iff$$
$$t_{ij} = \bar{t}_{ij}\left[t_{ij}^{-\theta} \times P_i^{-\theta} \times \Pi_j^{-\theta}\right]^{-\lambda} \iff$$
$$t_{ij} = \bar{t}_{ij}^{\frac{1}{1+\theta\lambda}} \times P_i^{-\frac{\theta\lambda}{1+\theta\lambda}} \times \Pi_j^{-\frac{\theta\lambda}{1+\theta\lambda}}$$

For both models, we have the equilibrium conditions:

$$\gamma_i = \sum_j X_{ij}$$

$$\delta_i = \sum_j X_{ji}$$

Starting with the first equilibrium conditions, we substitute in for our gravity model of flows and solve for market access term Π_i:

$$\begin{aligned}
\gamma_i &= \sum_j X_{ij} \iff \\
\gamma_i &= \sum_j \tau_{ij}^{-\theta} \times \frac{\gamma_i}{\Pi_i^{-\theta}} \times \frac{\delta_j}{P_j^{-\theta}} \iff \\
\gamma_i &= \sum_j \left(I - [t_{ij}]^{-\theta}\right)^{-1} \times \frac{\gamma_i}{\Pi_i^{-\theta}} \times \frac{\delta_j}{P_j^{-\theta}} \iff \\
\Pi_i^{-\theta} &= \sum_j \left(I - [t_{ij}]^{-\theta}\right)^{-1} \times \frac{\delta_j}{P_j^{-\theta}} \iff \\
\left(I - [t_{ij}]^{-\theta}\right) \Pi_i^{-\theta} &= \frac{\delta_i}{P_i^{-\theta}} \iff \\
\Pi_i^{-\theta} &= \frac{\delta_i}{P_i^{-\theta}} + \sum_j t_{ij}^{-\theta} \Pi_j^{-\theta}
\end{aligned}$$

For the second equilibrium condition, we do the same, but solving for market access term P_i:

$$\begin{aligned}
\delta_i &= \sum_j X_{ji} \iff \\
\delta_j &= \sum_j \tau_{ji}^{-\theta} \times \frac{\gamma_j}{\Pi_j^{-\theta}} \times \frac{\delta_i}{P_i^{-\theta}} \iff \\
P_i^{-\theta} &= \sum_j \tau_{ji}^{-\theta} \times \frac{\gamma_j}{\Pi_j^{-\theta}} \iff \\
P_i^{-\theta} &= \sum_j \left(I - [t_{ji}]^{-\theta}\right)^{-1} \times \frac{\gamma_j}{\Pi_j^{-\theta}} \iff \\
\left(I - [t_{ji}]^{-\theta}\right)^{-1} P_i^{-\theta} &= \frac{\gamma_i}{\Pi_i^{-\theta}} \iff \\
P_i^{-\theta} &= \frac{\gamma_i}{\Pi_i^{-\theta}} + \sum_j t_{ji}^{-\theta} P_j^{-\theta}
\end{aligned}$$

Expressed in changes, these two equilibrium conditions become

$$\hat{\Pi}_i^{-\theta} = \left(\frac{\frac{\delta_i}{P_i^{-\theta}}}{\frac{\delta_i}{P_i^{-\theta}} + \sum_j t_{ij}^{-\theta}\Pi_j^{-\theta}}\right) \frac{\hat{\delta}_i}{\hat{P}_i^{-\theta}} + \sum_j \left(\frac{t_{ij}^{-\theta}\Pi_j^{-\theta}}{\frac{\delta_i}{P_i^{-\theta}} + \sum_j t_{ij}^{-\theta}\Pi_j^{-\theta}}\right) \hat{t}_{ij}^{-\theta}\hat{\Pi}_j^{-\theta}$$

and:

$$\hat{P}_i^{-\theta} = \left(\frac{\frac{\gamma_i}{\Pi_i^{-\theta}}}{\frac{\gamma_i}{\Pi_i^{-\theta}} + \sum_j t_{ji}^{-\theta}P_j^{-\theta}}\right) \frac{\hat{\gamma}_i}{\hat{\Pi}_i^{-\theta}} + \sum_j \left(\frac{t_{ji}^{-\theta}P_j^{-\theta}}{\frac{\gamma_i}{\Pi_i^{-\theta}} + \sum_j t_{ji}^{-\theta}P_j^{-\theta}}\right) \hat{t}_{ji}^{-\theta}\hat{P}_j^{-\theta}$$

We multiply both the numerator and denominator by their appropriate market access term so that we can substitute in our expression for traffic, resulting in:

$$\hat{\Pi}_i^{-\theta} = \left(\frac{\delta_i}{\delta_i + \sum_j t_{ij}^{-\theta} P_i^{-\theta} \Pi_j^{-\theta}}\right) \frac{\hat{\delta}_i}{\hat{P}_i^{-\theta}} + \sum_j \left(\frac{t_{ij}^{-\theta} P_i^{-\theta} \Pi_j^{-\theta}}{\delta_i C + \sum_j t_{ij}^{-\theta} P_i^{-\theta} \Pi_j^{-\theta}}\right) \hat{t}_{ij}^{-\theta} \hat{\Pi}_j^{-\theta} \iff$$

$$\hat{\Pi}_i^{-\theta} = \left(\frac{\delta_i}{\delta_i + \sum_j \Xi_{ij}}\right) \frac{\hat{\delta}_i}{\hat{P}_i^{-\theta}} + \sum_j \left(\frac{\Xi_{ij}}{\delta_i C + \sum_j \Xi_{ij}}\right) \hat{t}_{ij}^{-\theta} \hat{\Pi}_j^{-\theta}$$

and:

$$\hat{P}_i^{-\theta} = \left(\frac{\gamma_i}{\gamma_i + \sum_j t_{ji}^{-\theta} \Pi_i^{-\theta} P_j^{-\theta}}\right) \frac{\hat{\gamma}_i}{\hat{\Pi}_i^{-\theta}} + \sum_j \left(\frac{t_{ji}^{-\theta} \Pi_i^{-\theta} P_j^{-\theta}}{\gamma_i C + \sum_j t_{ji}^{-\theta} \Pi_i^{-\theta} P_j^{-\theta}}\right) \hat{t}_{ji}^{-\theta} \hat{P}_j^{-\theta} \iff$$

$$\hat{P}_i^{-\theta} = \left(\frac{\gamma_i}{\gamma_i + \sum_j \Xi_{ji}}\right) \frac{\hat{\gamma}_i}{\hat{\Pi}_i^{-\theta}} + \sum_j \left(\frac{\Xi_{ji}}{\gamma_i + \sum_j \Xi_{ji}}\right) \hat{t}_{ji}^{-\theta} \hat{P}_j^{-\theta}$$

Finally, substituting in our expression for iceberg trade costs along a link,

$$\hat{t}_{ij} = \hat{\bar{t}}_{ij}^{\frac{1}{1+\theta\lambda}} \times \hat{P}_i^{-\frac{\theta\lambda}{1+\theta\lambda}} \times \hat{\Pi}_j^{-\frac{\theta\lambda}{1+\theta\lambda}}$$

and multiplying each equation by the other market access term, we obtain the following for the two equilibrium conditions:

$$\hat{\Pi}_i^{-\theta} \hat{P}_i^{-\theta} = \left(\frac{\delta_i}{\delta_i + \sum_j \Xi_{ij}}\right) \hat{\delta}_i + \sum_j \left(\frac{\Xi_{ij}}{\delta_i + \sum_j \Xi_{ij}}\right) \hat{\bar{t}}_{ij}^{-\frac{\theta}{1+\theta\lambda}} \hat{P}_i^{-\frac{\theta}{1+\theta\lambda}} \hat{\Pi}_j^{-\frac{\theta}{1+\theta\lambda}}$$

$$\hat{\Pi}_i^{-\theta} \hat{P}_i^{-\theta} = \left(\frac{\gamma_i}{\gamma_i + \sum_j \Xi_{ji}}\right) \hat{\gamma}_i + \sum_j \left(\frac{\Xi_{ji}}{\gamma_i + \sum_j \Xi_{ji}}\right) \hat{\bar{t}}_{ji}^{-\frac{\theta}{1+\theta\lambda}} \hat{P}_j^{-\frac{\theta}{1+\theta\lambda}} \hat{\Pi}_i^{-\frac{\theta}{1+\theta\lambda}}$$

Now that we have defined the counterfactual equations generally, we turn to the specific cases of the trade and traffic models. For the trade model, we have the following definitions for the fixed effects:

$$\delta_i = E_i$$

$$\gamma_i = Y_i$$

We also derive the following for the price indices:

$$P_i = \frac{w_i u_i}{W} \iff$$
$$P_i = Y_i \bar{u}_i L_i^{\beta-1} W^{-1} \implies$$
$$\hat{P}_i = \hat{y}_i \hat{l}_i^{\beta-1} \hat{W}^{-1}$$

and

$$\Pi_i = A_i L_i Y_i^{-\frac{\theta+1}{\theta}} \iff$$
$$\Pi_i = \bar{A}_i L_i^{\alpha+1} Y_i^{-\frac{\theta+1}{\theta}} \implies$$
$$\hat{\Pi}_i = \hat{l}_i^{\alpha+1} \hat{y}_i^{-\frac{\theta+1}{\theta}}$$

Substituting into the equilibrium conditions, we get:

$$
\begin{aligned}
\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_i\hat{l}_i^{\beta-1}\hat{W}^{-1}\right)^{-\theta} &= \left(\frac{E_i}{E_i+\sum_j \Xi_{ij}}\right)\hat{y}_i + \\
&\sum_j \left(\frac{\Xi_{ij}}{E_i+\sum_j \Xi_{ij}}\right)\hat{t}_{ij}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{y}_i\hat{l}_i^{\beta-1}\hat{W}^{-1}\right)^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_j^{\alpha+1}\hat{y}_j^{-\frac{\theta+1}{\theta}}\right)^{-\frac{\theta}{1+\theta\lambda}} \iff \\
\hat{y}_i\hat{l}_i^{-\theta(\alpha+\beta)}\left(\hat{y}_i\hat{l}_i^{\beta-1}\right)^{\frac{\theta}{1+\theta\lambda}} &= \left(\frac{E_i}{E_i+\sum_j \Xi_{ij}}\right)\hat{W}^{-\theta}\hat{y}_i\left(\hat{y}_i\hat{l}_i^{\beta-1}\right)^{\frac{\theta}{1+\theta\lambda}} + \\
&\hat{W}^{\frac{\theta}{1+\theta\lambda}-\theta}\sum_j \left(\frac{\Xi_{ij}}{E_i+\sum_j \Xi_{ij}}\right)\hat{y}_i\hat{t}_{ij}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_j^{\alpha+1}\hat{y}_j^{-\frac{\theta+1}{\theta}}\right)^{-\frac{\theta}{1+\theta\lambda}} \iff \\
\hat{y}_i^{\frac{1+\theta\lambda+\theta}{1+\theta\lambda}}\hat{l}_i^{-\frac{\theta(1+\alpha+\theta\lambda(\beta+\alpha))}{1+\theta\lambda}} &= \hat{\chi}\left(\frac{E_i}{E_i+\sum_k \Xi_{ik}}\right)\hat{y}_i^{\frac{1+\theta\lambda+\theta}{1+\theta\lambda}}\hat{l}_i^{\frac{\theta(\beta-1)}{1+\theta\lambda}} + \hat{\chi}^{\frac{\theta\lambda}{1+\theta\lambda}}\sum_j \left(\frac{\Xi_{ij}}{E_i+\sum_k \Xi_{ik}}\right)\hat{t}_{ij}^{-\frac{\theta}{1+\theta\lambda}}\hat{y}_j^{\frac{1+\theta}{1+\theta\lambda}}\hat{l}_j^{-\frac{\theta(1+\alpha)}{1+\theta\lambda}},
\end{aligned}
$$

as in equation (36) and:

$$
\begin{aligned}
\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_i\hat{l}_i^{\beta-1}\hat{W}^{-1}\right)^{-\theta} &= \left(\frac{Y_i}{Y_i+\sum_j \Xi_{ji}}\right)\hat{y}_i + \\
&\sum_j \left(\frac{\Xi_{ji}}{Y_i+\sum_j \Xi_{ji}}\right)\hat{t}_{ji}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{y}_j\hat{l}_j^{\beta-1}\hat{W}^{-1}\right)^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\frac{\theta}{1+\theta\lambda}} \iff \\
\hat{y}_i\hat{l}_i^{-\theta(\alpha+\beta)}\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{\frac{\theta}{1+\theta\lambda}} &= \left(\frac{Y_i}{Y_i+\sum_j \Xi_{ji}}\right)\hat{W}^{\theta}\hat{y}_i\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{\frac{\theta}{1+\theta\lambda}} + \\
&\hat{W}^{\frac{\theta}{1+\theta\lambda}-\theta}\sum_j \left(\frac{\Xi_{ji}}{Y_i+\sum_j \Xi_{ji}}\right)\hat{t}_{ji}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{y}_j\hat{l}_j^{\beta-1}\hat{W}^{-1}\right)^{-\frac{\theta}{1+\theta\lambda}} \iff \\
\hat{y}_i^{\frac{-\theta(1-\lambda)}{1+\theta\lambda}}\hat{l}_i^{\frac{\theta(1-\beta-\theta\lambda(\alpha+\beta))}{1+\theta\lambda}} &= \hat{\chi}\left(\frac{Y_i}{Y_i+\sum_k \Xi_{ki}}\right)\hat{y}_i^{\frac{-\theta(1-\lambda)}{1+\theta\lambda}}\hat{l}_i^{\frac{\theta(\alpha+1)}{1+\theta\lambda}} + \hat{\chi}^{\frac{\theta\lambda}{1+\theta\lambda}}\sum_{j=1}^{N}\left(\frac{\Xi_{ji}}{Y_i+\sum_k \Xi_{ki}}\right)\hat{y}_j^{-\frac{\theta}{1+\theta\lambda}}\hat{l}_j^{\frac{\theta(1-\beta)}{1+\theta\lambda}}
\end{aligned}
$$

as in equation (37).

For the commuting model, we define the following for fixed effects:

$$\gamma_i = L_i^R$$

$$\delta_i = L_i^F$$

We also derive the following for the price indices:

$$
\begin{aligned}
P_i &= A_i\left(L_i^F\right)^{-\frac{1}{\theta}}W^{-1} \iff \\
P_i &= \bar{A}_i\left(L_i^F\right)^{\alpha-\frac{1}{\theta}}W^{-1} \implies \\
\hat{P}_i &= \left(\hat{l}_i^F\right)^{\alpha-\frac{1}{\theta}}\hat{W}^{-1}
\end{aligned}
$$

and:

$$\Pi_i = u_i \left(L_i^R\right)^{-\frac{1}{\theta}} \iff$$
$$\Pi_i = \bar{u}_i \left(L_i^R\right)^{\beta-\frac{1}{\theta}} \implies$$
$$\hat{\Pi}_i = \left(\hat{l}_i^R\right)^{\beta-\frac{1}{\theta}}$$

Substituting these into our generalized equilibrium conditions, we get:

$$\left(\hat{l}_i^R\right)^{1-\beta\theta}\left(\hat{l}_i^F\right)^{1-\alpha\theta}\hat{W}^\theta = \left(\frac{L_i^F}{L_i^F+\sum_j \Xi_{ij}}\right)\hat{l}_i^F + \sum_j \left(\frac{\Xi_{ij}}{L_i^F+\sum_j \Xi_{ij}}\right)\hat{t}_{ij}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_i^F\right)^{\frac{1-\alpha\theta}{1+\theta\lambda}}\hat{W}^{\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_j^R\right)^{\frac{1-\beta\theta}{1+\theta\lambda}} \iff$$
$$\left(\hat{l}_i^R\right)^{1-\beta\theta}\left(\hat{l}_i^F\right)^{1-\alpha\theta-\frac{1-\alpha\theta}{1+\theta\lambda}} = \left(\frac{L_i^F}{L_i^F+\sum_j \Xi_{ij}}\right)\hat{W}^{-\theta}\left(\hat{l}_i^F\right)^{1-\frac{1-\alpha\theta}{1+\theta\lambda}} + \hat{W}^{\frac{\theta}{1+\theta\lambda}-\theta}\sum_j \left(\frac{\Xi_{ij}}{L_i^F+\sum_j \Xi_{ij}}\right)\hat{t}_{ij}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_j^R\right)^{\frac{1-\beta\theta}{1+\theta\lambda}} \iff$$
$$\left(\hat{l}_i^R\right)^{1-\beta\theta}\left(\hat{l}_i^F\right)^{\frac{\theta\lambda(1-\alpha\theta)}{1+\theta\lambda}} = \hat{\chi}\left(\frac{L_i^F}{L_i^F+\sum_j \Xi_{ij}}\right)\left(\hat{l}_i^F\right)^{\frac{\theta(\alpha+\lambda)}{1+\theta\lambda}} + \hat{\chi}^{\frac{\theta\lambda}{1+\theta\lambda}}\sum_j \left(\frac{\Xi_{ij}}{L_i^F+\sum_j \Xi_{ij}}\right)\hat{t}_{ij}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_j^R\right)^{\frac{1-\beta\theta}{1+\theta\lambda}},$$

as in equation (38) and:

$$\left(\hat{l}_i^R\right)^{1-\beta\theta}\left(\hat{l}_i^F\right)^{1-\alpha\theta}\hat{W}^\theta = \left(\frac{L_i^R}{L_i^R+\sum_j \Xi_{ji}}\right)\hat{l}_i^R + \sum_j \left(\frac{\Xi_{ji}}{L_i^R+\sum_j \Xi_{ji}}\right)\hat{t}_{ji}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_j^F\right)^{\frac{1-\alpha\theta}{1+\theta\lambda}}\hat{W}^{\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_i^R\right)^{\frac{1-\beta\theta}{1+\theta\lambda}} \iff$$
$$\left(\hat{l}_i^R\right)^{1-\beta\theta-\frac{1-\beta\theta}{1+\theta\lambda}}\left(\hat{l}_i^F\right)^{1-\alpha\theta} = \left(\frac{L_i^R}{L_i^R+\sum_j \Xi_{ji}}\right)\hat{W}^{-\theta}\left(\hat{l}_i^R\right)^{1-\frac{1-\beta\theta}{1+\theta\lambda}} + \hat{W}^{\frac{\theta}{1+\theta\lambda}-\theta}\sum_j \left(\frac{\Xi_{ji}}{L_i^R+\sum_j \Xi_{ji}}\right)\hat{t}_{ji}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_j^F\right)^{\frac{1-\alpha\theta}{1+\theta\lambda}} \iff$$
$$\left(\hat{l}_i^R\right)^{\frac{\theta\lambda(1-\beta\theta)}{1+\theta\lambda}}\left(\hat{l}_i^F\right)^{1-\alpha\theta} = \hat{\chi}^\theta\left(\frac{L_i^R}{L_i^R+\sum_j \Xi_{ji}}\right)\left(\hat{l}_i^R\right)^{\frac{\theta(\beta+\lambda)}{1+\theta\lambda}} + \hat{\chi}^{\frac{\theta\lambda}{1+\theta\lambda}}\sum_j \left(\frac{\Xi_{ji}}{L_i^R+\sum_j \Xi_{ji}}\right)\hat{t}_{ji}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_j^F\right)^{\frac{1-\alpha\theta}{1+\theta\lambda}},$$

as in equation (39).

D Online Appendix: Extensions

In this section, we present two model extensions mentioned in the main text.

D.1 Additive Transportation Costs

In this subsection, we describe an alternative framework where transportation costs are additive across segments (rather than multiplicative, as assumed in the main text). Let L_{ij} be the worker hours employed in the production and shipment of goods from i to j, which is split into the hours workers spend producing the good, $L_{ij}^{produce}$ and hours the workers spend shipping the good L_{ij}^{ship}:

$$L_{ij} = L_{ij}^{produce} + L_{ij}^{ship}$$

The total number of goods being sent from i to j is:

$$Q_{ij} = A_i L_{ij}^{produce}$$

Let t_{ij} be the expected travel time from i to j (see below). Suppose that each unit of good requires a separate truck so that the total amount of labor hours used in shipping goods is:

$$L_{ij}^{ship} = Q_{ij} t_{ij} = A_i L_{ij}^{produce} t_{ij}.$$

Hence, to produce each unit of a good to send from i to j requires $\frac{(1+t_{ij})}{A_i}$ units of labor and costs $p_{ij} = (1+t_{ij}) \frac{w_i}{A_i}$. If we define $\tau_{ij} \equiv 1 + t_{ij}$, the model is identical to the economic geography presented in the main text.

Now we calculate the expected cost. Let μ_{ij} the direct travel time between i and j. The total aggregate travel time from i to j on path p of length K is:

$$\tilde{t}_{ij}(p) = \sum_{k=1}^{K} \mu_{p_{k-1},p_k}$$

Suppose each worker $\nu \in [1, L_i]$ is heterogeneous in her preferences of routes so that she chooses the path p to minimize:

$$\hat{t}_{ij}(\nu) = \min_{K \geq 0, p \in P_{ij}^K} \tilde{t}_{ij}(p) + \varepsilon_{ij}(p, \nu),$$

where $\varepsilon_{ij}(p, \nu)$ is Gumbel distributed with shape parameter θ.

The expected trade cost can then be written as:

$$t_{ij} = -\frac{1}{\theta} \ln \sum_{K \geq 0, p \in P_{ij}^K} \exp\left(\tilde{t}_{ij}(p)\right)^{-\theta} \iff$$

$$t_{ij} = -\frac{1}{\theta} \ln \sum_{K \geq 0, p \in P_{ij}^K} \exp\left(\sum_{k=1}^{K} \mu_{p_{k-1},p_k}\right)^{-\theta} \iff$$

$$t_{ij} = -\frac{1}{\theta} \ln \sum_{K=0}^{\infty} \sum_{p \in P_{ij}^K} \left(\prod_{k=1}^{K} \exp\left(-\theta \mu_{p_{k-1},p_k}\right)\right) \iff$$

$$t_{ij} = -\frac{1}{\theta} \ln \sum_{K=0}^{\infty} \left(\sum_{k_1=1}^{N} \sum_{k_2=1}^{N} .. \sum_{k_{K-1}=1}^{N} \left(a_{i,k_1} \times a_{k_1,k_2} \times ... \times a_{k_{K-2},k_{K-1}} \times a_{k_{K-1},j}\right)\right) \iff,$$

$$t_{ij} = -\frac{1}{\theta} \ln \sum_{K=0}^{\infty} A_{ij}^K$$

where $a_{ij} \equiv \exp(-\theta \mu_{ij})$, $\mathbf{A} \equiv [a_{ij}]$ and $\mathbf{A}^K \equiv \left[A_{ij}^K\right]$. Define $\mathbf{B} \equiv (\mathbf{I} - \mathbf{A})^{-1} = \sum_{K=0}^{\infty} \boldsymbol{A}^K$. We then have:

$$t_{ij} = -\frac{1}{\theta} \ln b_{ij}.$$

Hence, the iceberg transportation cost with additive transportation costs can be written as:

$$\tau'_{ij} = 1 + \ln\left(b_{ij}^{-\frac{1}{\theta}}\right) \approx b_{ij}^{-\frac{1}{\theta}},$$

i.e. it is approximately equal to the iceberg transportation cost defined in equation (21).

D.2 Nested Route Choice

In this subsection, we present an alternative economic geography model where agents first choose from which location to source the good and then choose on which route to ship the good. Suppose that each location $i \in \mathcal{N}$ is endowed with a constant returns to scale technology for producing and shipping each good $\nu \in [0, 1]$

to each destination $j \in \mathcal{N}$ along each route $r \in \Re_{ij}$, which under perfect competition yields the following price of good $\nu \in [0,1]$ in destination $j \in \mathcal{N}$ from origin $i \in \mathcal{N}$ along route $r \in \Re_{ij}$:

$$p_{ij,r}(\nu) = \frac{w_i}{\varepsilon_i(\nu)} \times \frac{\prod_{k=1}^{K} t_{r_{k-1},r_k}}{\epsilon_{ij,r}(\nu)},$$

where $\varepsilon_i(\nu)$ is an independently and identically Frechet distributed across goods distributed with scale parameter $1/A_i$ and shape parameter θ_g and $\epsilon_{ij,r}(\nu)$ is independently and identically Frechet distributed across routes with a scale parameter equal to $\Gamma\left(\frac{\theta_r-1}{\theta_r}\right)^{-1}$ (which is done without loss of generality and for convenience alone) and shape parameter θ_r. The timing is as follows: first, individuals observe $\varepsilon_i(\nu)$ and choose a source to purchase the good; second, individuals observe $\epsilon_{ij,r}(\nu)$ and choose the route to ship the good. (For simplicity, individuals are not permitted to alter their decision of where to source the good once $\epsilon_{ij,r}(\nu)$ is revealed).

To solve the model, we proceed by backwards induction. Conditional on choosing to source good ν from location i, a consumer in location j will choose a route from i to j to minimize the shipping cost incurred, so that the probability she selects a route $r \in \Re_{ij}$ is:

$$\pi_{ij,r|i} = \frac{\left(\prod_{k=1}^{K} t_{r_{k-1},r_k}\right)^{\theta_r}}{\sum_{r' \in \Re_{ij}} \left(\prod_{k=1}^{K} t_{r'_{k-1},r'_k}\right)^{\theta_r}}$$

and the expected cost she incurs in shipping a good from i to j is:

$$\tau_{ij} = \left(\sum_{r' \in \Re_{ij}} \left(\prod_{k=1}^{K} t_{r'_{k-1},r'_k}\right)^{-\theta_r}\right)^{-\frac{1}{\theta_r}} \tag{D.1}$$

Apart from the different θ, equation (D.1) is equivalent to equation (3) and so a similar be expressed equivalently as in equation (21):

$$\tau_{ij} = b_{ij}^{-\frac{1}{\theta_r}}, \tag{D.2}$$

where b_{ij} is the (i,j) element of the matrix $\mathbf{B} = (\mathbf{I} - \mathbf{A})^{-1}$ and $\mathbf{A} \equiv [a_{ij}] = \left[t_{ij}^{-\theta_r}\right]$.

Knowing that this is the expected cost she will incur, a consumer in location j will first choose the location i to source the good ν from in order to minimize the expected total cost, so that the probability she sources from $i \in \mathcal{N}$ is:

$$\pi_{ij} = \frac{\tau_{ij}^{-\theta_g}(w_i/A_i)^{-\theta_g}}{\sum_{k \in \mathcal{N}} \tau_{kj}^{-\theta_g}(w_k/A_k)^{-\theta_g}} \tag{D.3}$$

and the total value of trade from i to j can be written as:

$$X_{ij} = \frac{\tau_{ij}^{-\theta_g}(w_i/A_i)^{-\theta_g}}{\sum_{k \in \mathcal{N}} \tau_{kj}^{-\theta_g}(w_k/A_k)^{-\theta_g}} E_j. \tag{D.4}$$

It is immediately evident that when $\theta_r = \theta_g = \theta$, equation (D.2) is identical to equation (21) and equation (D.4) is identical to equation (2), i.e. the model here becomes isomorphic to the one presented in the main text. More generally, (and to get a sense of where the tractability is lost when $\theta_r \neq \theta_g$), combining equations (D.2) and (D.4) yields:

$$X_{ij} = \frac{b_{ij}^{\left(\frac{\theta_g}{\theta_r}\right)}(w_i/A_i)^{-\theta_g}}{\sum_{k \in \mathcal{N}} b_{kj}^{\left(\frac{\theta_g}{\theta_r}\right)}(w_k/A_k)^{-\theta_g}} E_j,$$

so the bilateral trade flows are functions of elements of the Leontief inverse raised to a power (rather than the elements of the Leontief inverse themselves).

E Online Appendix: Algorithm for Conducting Counterfactuals

The algorithm we use to find the equilibrium in our counterfactual simulations consists of an outer loop, where we guess a $\hat{\chi}$, and an inner loop, where, given a $\hat{\chi}$, we solve for vectors of $\{\hat{y}_i\}$ and $\left\{\hat{l}_i\right\}$ in the economic geography case and vectors of $\left\{\hat{l}_i^R\right\}$ and $\left\{\hat{l}_i^F\right\}$ in the commuting case for which the system is equal up to scale. We can see that for equilibrium equations (36) and (37) (the economic geography model) and equilibrium equations (38) and (39) (the commuting model), given a $\hat{\chi}$, the system forms a system of non-linear equations in $\left(\{\hat{y}_i\}, \left\{\hat{l}_i\right\}\right)$ and $\left(\left\{\hat{l}_i^R\right\}, \left\{\hat{l}_i^F\right\}\right)$, respectively. At the same time, the term on the left-hand side of each of the equilibrium conditions is a log-linear combination of the endogenous variables. Because of these similarities, we use a very similar algorithm to solve both models. For concision, we will describe the algorithm in detail in terms of the endogenous variables of the economic geography model.

Let's start with a detailed description of the inner loop. Given an initial guess of $\hat{\chi}_{(0)} = 1$, we plug in an initial guess of the endogenous variables $\{\hat{y}_i\}_{(0)} = 1$ and $\left\{\hat{l}_i\right\}_{(0)} = 1$ into our equilibrium conditions (38) and (39). To help us update our guess, we define the following:

$$\hat{x}_{1,i} \equiv \hat{y}_i^{\frac{1+\theta\lambda+\theta}{1+\theta\lambda}} \hat{l}_i^{-\theta\left(\frac{1+\alpha+\theta\lambda(\alpha+\beta)}{1+\theta\lambda}\right)}$$

$$\hat{x}_{2,i} \equiv \hat{y}_i^{-\theta\left(\frac{1-\lambda}{1+\theta\lambda}\right)} \hat{l}_i^{-\theta\left(\frac{-1+\beta+\theta\lambda(\alpha+\beta)}{1+\theta\lambda}\right)}$$

so that:

$$\begin{pmatrix} \ln \hat{x}_{1,i} \\ \ln \hat{x}_{2,i} \end{pmatrix} = \begin{pmatrix} \frac{1+\theta\lambda+\theta}{1+\theta\lambda} & -\theta\left(\frac{1+\alpha+\theta\lambda(\alpha+\beta)}{1+\theta\lambda}\right) \\ -\theta\left(\frac{1-\lambda}{1+\theta\lambda}\right) & -\theta\left(\frac{-1+\beta+\theta\lambda(\alpha+\beta)}{1+\theta\lambda}\right) \end{pmatrix} \begin{pmatrix} \ln \hat{y}_i \\ \ln \hat{l}_i \end{pmatrix} \iff$$

$$\begin{pmatrix} \ln \hat{y}_i \\ \ln \hat{l}_i \end{pmatrix} = \begin{pmatrix} \frac{1+\theta\lambda+\theta}{1+\theta\lambda} & -\theta\left(\frac{1+\alpha+\theta\lambda(\alpha+\beta)}{1+\theta\lambda}\right) \\ -\theta\left(\frac{1-\lambda}{1+\theta\lambda}\right) & -\theta\left(\frac{-1+\beta+\theta\lambda(\alpha+\beta)}{1+\theta\lambda}\right) \end{pmatrix}^{-1} \begin{pmatrix} \ln \hat{x}_{1,i} \\ \ln \hat{x}_{2,i} \end{pmatrix}$$

By this definition, for a guess of $\left(\{\hat{y}_i\}_{(0)}, \left\{\hat{l}_i\right\}_{(0)}\right)$, the equilibrium conditions yield a set of vectors $\left(\{\hat{x}_{1,i}\}_{(1)}, \{\hat{x}_{2,i}\}_{(1)}\right)$, which by the log-linear transformation defined above, imply an updated guess of $\left(\{\hat{y}_i\}_{(1)}, \left\{\hat{l}_i\right\}_{(1)}\right)$. On each iteration, we rescale the vectors $\left(\{\hat{x}_{1,i}\}_{(1)}, \{\hat{x}_{2,i}\}_{(1)}\right)$ such that the second-period income and labor distribution still sum to 1, and update our guess of $\left(\{\hat{y}_i\}_{(0)}, \left\{\hat{l}_i\right\}_{(0)}\right)$ towards $\left(\{\hat{y}_i\}_{(1)}, \left\{\hat{l}_i\right\}_{(1)}\right)$, We iterate through procedure this until the equilibrium conditions are solved up to scale. Therefore, the inner loop returns a $\left(\{\hat{y}_i\}_{(0)}, \left\{\hat{l}_i\right\}_{(0)}\right)$ for which:

$$\lambda_{1,i} \left(\hat{x}_{1,i}\right)_{(1)} = \hat{\chi}_{(0)} \left(\frac{E_i}{E_i + \sum_k \Xi_{ik}}\right) \left(\hat{y}_i\right)_{(0)}^{\frac{1+\theta\lambda+\theta}{1+\theta\lambda}} \left(\hat{l}_i\right)_{(0)}^{\frac{\theta(\beta-1)}{1+\theta\lambda}} + \hat{\chi}_{(0)}^{\frac{\theta\lambda}{1+\theta\lambda}} \sum_j \left(\frac{\Xi_{ij}}{E_i + \sum_k \Xi_{ik}}\right) \hat{\bar{t}}_{ij}^{-\frac{\theta}{1+\theta\lambda}} \left(\hat{y}_j\right)_{(0)}^{\frac{1+\theta}{1+\theta\lambda}} \left(\hat{l}_j\right)_{(0)}^{-\frac{\theta(1+\alpha)}{1+\theta\lambda}}$$

$$\lambda_{2,i}\left(\hat{x}_{2,i}\right)_{(1)} = \hat{\chi}_{(0)}\left(\frac{Y_i}{Y_i+\sum_k \Xi_{ki}}\right)\left(\hat{y}_i\right)_{(0)}^{\frac{-\theta(1-\lambda)}{1+\theta\lambda}}\left(\hat{l}_i\right)_{(0)}^{\frac{\theta(\alpha+1)}{1+\theta\lambda}} + \hat{\chi}_{(0)}^{\frac{\theta\lambda}{1+\theta\lambda}}\sum_{j=1}^{N}\left(\frac{\Xi_{ji}}{Y_i+\sum_k \Xi_{ki}}\right)\left(\hat{y}_j\right)_{(0)}^{-\frac{\theta}{1+\theta\lambda}}\left(\hat{l}_j\right)_{(0)}^{\frac{\theta(1-\beta)}{1+\theta\lambda}}$$

To solve the equilibrium conditions, we'd like to have a $\hat{\chi}$ for which $\{\lambda_{1,i}\} = 1$ and $\{\lambda_{2,i}\} = 1$. Therefore, in the outer loop we implement a version of the fmincon function, a non-linear constrained minimization function built into MATLAB, which finds a $\hat{\chi}_{(1)}$ which minimizes the matrix norm of $[\{\ln\lambda_{1,i}\}\ \{\ln\lambda_{2,i}\}]$. We use the output of fmincon to update our guess of $\hat{\chi}$, and we iterate on this outer loop until $\hat{\chi}_{(0)}$ converges with $\hat{\chi}_{(1)}$.

F Online Appendix: Data Construction

In this section, we provide details on the construction of the data used in the paper.

F.1 The U.S. Highway System Network

Our version of the interstate highway system consists of road segments, which are parts of an interstate which pass through a county. Each segment has data on mileage, traffic flows, and through lanes over the entirety of the segment, as well as over subsections of the segment which fall into different road type classifications under the Federal Highway Administration's (FHA) Highway Economic Requirements System (HERS). To generate this dataset, we combine GIS data on the interstate highway network released by the FHA, elevation data from NASA's Shuttle Radar Topography Mission (SRTM), population data on and geographies of urban areas, Commodity Flow Survey (CFS) Areas, and Census-Based Statistical Areas (CBSA) released by the Census Bureau and sourced through IPUMS NH-GIS, and trade information from the CFS.

F.1.1 Interstate Highway System & Traffic Data

Since 2011, the FHA has released shapefiles of the national road network, which are linked to data collected from the Highway Performance Monitoring System (HPMS). Critically, this shapefile contains information on the state and county a road segment is in, the roads' Department of Transportation functional system classification, a road's name or route number, whether it is in an urban or rural area and which urban area it is in, the mileage of a segment, the average annual daily traffic (AADT) that goes over it, and how many lanes the segment has.

To generate our version of the interstate highway system, we begin with the HPMS release from 2012, trim it to only those roads within the contiguous United States, and remove all roads that are not classified as Interstates, resulting a road network with 334,040 segments. Beyond this, we resolve several data quality issues within the 2012 HPMS release. The most visible of these issues is the road reports from West Virginia, which are missing large sections of several interstate highways. To resolve these, we replace the reported data on West Virginia interstates from 2012 with that from 2013. Additionally, because the national HPMS dataset is sourced from reports prepared by state departments of transportation, there are some discrepancies in how roads are labeled. For instance, for Interstate Route 10, one department of transportation might code it as 10, another as I10, another as I-10, and another might use an entirely different identification system altogether. In order to consistently code each road segment by its integer route number (i.e. Interstate 10 as "10"), it was necessary to recode interstate segments in Arizona, California, Nevada, and Rhode Island, with reference to Google Maps. Finally, the HPMS has some road segments which are classified as interstates but are not coded with a route number at all. These "zeroes" come from three states: Alabama, Maine, and New York. We delete those in Alabama, being exceptionally small in length (<0.1 miles), and those in Maine, which are short ramps, insignificant to the overall connectivity of the road network. The "zero" in New York is a transitional ramp between I-90 and I-87 near Albany and is recoded as part of I-90. The resulting dataset consists of 333,021 road segments.

Table F.1: Interstate Highway System: Cost of Adding a Lane-Mile

Road Type	Cost ($m)
Rural - Flat	1.923
Rural - Rolling	2.085
Rural - Mountainous	6.492
Small Urban	3.061
Small Urbanized	3.345
Large Urbanized	5.598
Major Urbanized	11.197

Source: Federal Highway Administration (2015)

F.1.2 Road Type Classifications

Using the cleaned version of the interstate highway system, we proceed to match each segment of the highway system to its underlying terrain. An elevation raster of the United States was sourced from DIVA-GIS, which gathers the underlying elevation data for the raster from the CGIAR's Center for Spatial Information SRTM data. Using this elevation raster and ArcGIS' 3D Analyst toolkit, we extract the average grade of each segment of the interstate highway system.

Each road segment is then matched with the population of the urban area it passes through based on its urban code in the HPMS data, which are the same codes that the Census Bureau uses to identify its urban areas and urbanized centers. Populations for urban areas are sourced from the 2010 Census Urban-Rural Classification, which was released in March 2012. The urban area codes in use in the HPMS data differ slightly from those in the census release, so a handful of urban areas in the HPMS data are recoded to match their updated codes in 2010 Census Urban-Rural Classification. Based on this terrain and population data, each segment was then classified into one of the seven HERS urban-rural road types below.

Urban road segments are classified based on population, per standards outlined by the FHA in its HPMS field manual Federal Highway Administration (2016). Rural road segments, which are all segments which pass through areas of less than 5,000 in population, are classified based on the average grade. The FHA offers only general guidance on how to classify roads by terrain. Based upon the guidance that Level terrain "generally includes short grades of no more than 2 percent" Federal Highway Administration (2016), all roads of grade below 2% were classified as Level, and based upon the maximum grade for Interstate Highways going over rolling terrain with a speed limit of 60 mph being set at 4% American Association of State Highway and Transportation Officials (2016), all roads of grades between 2% and 4% were classified as Rolling. The remainder of roads (those over 4% in grade) were classified as Mountainous.

For each section, a measure of its length is generated by subtracting the mileage marker of its endpoint from the mileage marker of its beginning point. Then, we generate a measure of vehicle miles traveled (VMT) by multiplying AADT by the length of the segment and a measure of lane-miles over the segment by multiplying through lanes (the total two-way lane width of a road) by the length of the segment. Each of these three measures—length, VMT, and lane-miles—is also interacted with the seven dummy variables that code road type.

F.1.3 Observed Network of the Interstate System

To simplify the geometry of the interstate network, we aggregate road segments based on their state, county, and route number, summing length, VMT, lane-miles, and all road type-interactions. This reduces the number of road segments from 333,021 to 1,761. Finally, we join segments within a radius of 3500 meters of each other together. This links together geometries which were either not precisely connected in the shapefile or were connected by shorter roads not coded as interstates in the HPMS dataset, and therefore removed in the initial data cleaning. This dataset forms the links of our interstate network.

For our network analysis of the interstate system, we choose to place nodes at every intersection between two different interstates and endpoint in the interstate highway system. This results in a set of 616 nodes, each of which we geocode with its latitude, longitude, distance from and name of nearest CBSA, and CFS area. We also identify adjacent nodes as nodes which can be reached from each another without passing through another node on the network.

F.1.4 Simplified Network of the Interstate System

In ArcMap, we set the nodes as origins and destinations in a symmetric origin-destination matrix; identify length, VMT, lane-miles and all road type-interactions as accumulation attributes (that is, fields that the ArcMap Network Analyst toolbox should integrate over as it calculates the least-cost route); and set length as the impedance, which is the field that the Network Analyst toolbox minimizes to identify the least-cost route. Solving this symmetric origin-destination matrix between 616 origins and 616 destination leads to 379,456 bilateral connections. Each of these bilateral connections contains data on the length of the route, total VMT over the route, total lane-miles over the route, and the length, total VMT, and total lane-miles over sections of the route which fall into each of the seven road types.

Using the CBSA information geocoded to the nodes, we code a node as being within its nearest CBSA if it is less than 3000 m away from the boundary of that CBSA. This is necessary to address the presence of nodes which just barely fall outside of a CBSA. The set of bilateral connections generated by solving the origin-destination matrix has several "redundant" connections because our approach to generating the nodes of the interstate network yields clusters of nodes near cities, with large ring roads or through which many interstate highways pass. To eliminate these redundancies, we consolidate all nodes coded to the same CBSA into one "CBSA node," coded with the average latitude and longitude of all nodes within that CBSA and its relevant CBSA and CFS area. The bilateral connections from this "CBSA node" to all other nodes in the interstate network contain the average distance units (length, VMT, lane-miles, and road type interactions) for each unique connection from that CBSA to another node. This process of consolidating nodes within CBSAs yields a simplified adjacency matrix, where the clusters of nodes around major cities due to ring roads are absorbed into one node. The 616 nodes and 379,456 links of the original OD matrix are reduced down to 228 nodes and 51,984 links, of which 704 are between adjacent nodes.

F.1.5 Estimated Cost of Improvements and Congestion

For the simplified network, we calculate the cost of adding an additional lane-mile along an link by identifying the share of each link that goes over each road type and using those shares as weights in a weighted average of the different cost figures for adding a lane-mile in each terrain type estimated by the Federal Highway Administration (2015). To identify congestion measures, we divide the total VMT along an link by the total lane-miles along that link. This gives us a measure of traffic per lane-mile over the course of the road.

F.1.6 Consistent Measures of Node Population and Income

To generate a consistent measure of population and income at each node, we sum the population of and average the median income of all cities within 25 miles of a node, conditional on a city within that radius not being closer to another node. We name each node (for readability) after the city with the largest population in the aforementioned 25 mile radius. Population and median income data come from a purchased dataset from USCitiesList.org. Although consistent, this way of measuring population naturally tends to understate populations for less densely populated areas.

F.1.7 Trade Flows: Observed and Imputed

Using the CFS area coded to each node, we link links with trucking flows aggregated to the origin-destination level from the 2012 CFS. CFS areas are generally larger than CBSA's, so to get more granular trade flows, we imputed commodity flows between CBSAs by assuming that, for CFS areas which consist of more than one CBSA, each CBSA receives and sends out a portion of flows proportional to its share of the CFS area's total GDP. Both the observed CFS area-CFS area flows and imputed CBSA-CBSA flows are included in the

Table F.2: Seattle Road Network: Cost of Adding a Lane-Mile

Road Type	Standard Cost ($m)	High Cost ($m)
Freeway/Interstate	11.197	46.691
Other Principal Arterial	8.252	31.988
Minor Arterial/Collector	5.614	31.988
Local*	5.614	N/A

*Local costs imputed as identical to Minor Arterial/Collector costs
Source: Federal Highway Administration (2015)

data. 9,801 of the 51,984 links are linked to CFS flows; 9,651 of those links can be further disaggregated into CBSA-CBSA flows.

F.2 The Seattle Road Network

Our version of Seattle's commuting network combines the road system reported in the Seattle-Tacoma-Bellevue CBSA in the FHA's 2016 HPMS release for the state of Washington with commuting flow data from the 2017 LEHD LODES release from the Census Bureau. We also, similarly, to the our version of the interstate highway network rely on elevation data from the SRTM and population data for urban areas and Census block groups released by the Census Bureau.

F.2.1 Local Road Data

We trim the data from the 2016 HPMS release for the state of Washington to cover all roads within the municipal boundaries of Seattle, creating a dataset of 9,188 road segments. This dataset contains information on a road segment's Department of Transportation functional system classification, which authority owns it, how many lanes it has, whether additional lanes could be easily added, and the AADT that flows over it, and whether it is a ramp or not.

There are a handful of roads for which one or several of these datapoints are blank, so we impute those based upon the features of the surrounding roads. For functional system classification and ownership, which are both categorical variables with an associated hierarchy, we fill in these blanks with the "highest" level of the hierarchy that a road comes into contact with; for example, if a road touches an interstate highway and a minor arterial, it is classified as belonging to the interstate functional classification, and if a road touches a road that is owned by the state and another that is owned by the county, that road is classified as being owned by the state. Generally, for lane width, we fill in blanks with the maximum lane width among roads that a segment comes into contact with, and for traffic flows, we fill in blanks with the mean traffic flows among roads that a segment comes into contact with. For a subset of roads which happens to have blank lane widths and traffic flows because they represent the other lane of a dual lane road way with that data, like a large highway or boulevard, we simply impute traffic flows and lane width from the its parallel counterpart. The only other exception to the aforementioned general rule is surface streets, where we impute that each surface street has width of two lanes and an AADT of 120. Finally, we used geoprocessing tools in ArcGIS to fix connectivity issues in the road data.

Taking the cleaned traffic data, we prepared the road network for network analysis by aggregating road segment-level data to what we call the "road section"-level—road sections being defined as a continuous segment of road not interrupted by an intersection with another road. This increases the number of road segments in our data to 17,261. We used ArcGIS tools to measure the length of each road section and code it to the Census urban area it belongs to. We further classified each road as being a high cost road to add a lane to or a not, based on whether the HPMS release says that additional lanes could be easily added to it. Using a road section's functional system classification and its high-cost classification, we coded each road section with the cost of adding a lane-mile to it, based on the costs estimated by the FHA Federal Highway Administration (2015).

Finally, for road sections which are missing posted speed limits, we filled in speed limits based on whether it is a ramp (ramps are given a speed limit of 30 mph) and who owns the road. Washington state has default statutory speed limits for city and town streets, county roads, and state highways Washington State Legislature (1965). Local roads are assigned a speed limit of 20 miles per hour, as per local Seattle traffic regulations. From these road type classifications, the length of the road, and the traffic and through-lane capacity data from the HPMS release, we created measures of VMT, lane-miles, improvement cost—generated by multiplying the cost of adding a lane-mile to a road by its length—and unimpeded travel time—generated by dividing the length of the road by its speed limit.

F.2.2 The Observed Road Network

We convert this local data into the observed road network by generating a set of nodes for network analysis, solving for the least-cost path between those nodes using ArcMap's Network Analyst tool, and paring the resulting set of bilateral paths down to only those between adjacent nodes.

To generate the nodes, we grid the city of Seattle into 224 1 km x 1 km parcels and set the center points of those parcels as nodes for network analysis. We restrict participation in the network analysis to only those nodes which are within a third of a kilometer of the road network; we find that this distance restriction does a good job of resolving the tradeoff between capturing the overall structure of Seattle's roads and limiting the nodes to those within a reasonable distance of the road system. Overall, 217 nodes participate in the network analysis.

Using the OD Matrix feature of the Network Analyst tool, we solve for the path which minimizes unimpeded travel time from each of the 217 nodes to all the other nodes. Along each path, we also sum over VMT, lane-miles, improvement cost, the number of intersections crossed, the number of turns taken, and the length traveled along arterial and local roads. While the network dataset has highly detailed data on all the roads in Seattle, this step means that we observe only those roads along which at least one least cost route between nodes travel. Solving this optimization problem yields a dataset with 47,089 bilateral connections between nodes. We define a node as adjacent to another node if it is in one of eight parcels which surrounds the other node's parcel and is not separated from the other node by a body of water, without being connected by a bridge.[39] There are 1,384 bilateral connections between adjacent nodes in our dataset.

F.2.3 Node Populations and Incomes

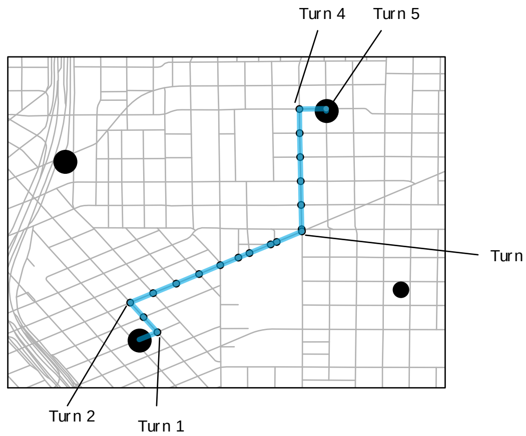
Since our nodes are not linked to any existing administrative dataset on populations and incomes, we need to generate population and income figure for each one. To do so, we use an ArcMap tool to identify the area of each intersection between a block group and a parcel and calculate the share of each parcel's area that comes from a particular block group. We also calculate the population density of each block group. Then, assuming that the population of each block group is uniformly distributed within that block group, we estimate the population density of each parcel by finding the weighted average of the block groups it overlaps with, where the weights are provided by the share of each parcel's area that comes from a block group. We calculate the per capita income of each parcel using a similar method. Further assuming that the residential and working population within each parcel is uniformly distributed, we calculate the total residential and working population of each parcel by multiplying the respective population density by its area.

F.2.4 Commuting Flows

We take commuting flow data from the LEHD LODES dataset. We narrow down the commuting flows, which originally are at the census block-to-census block level, to only those flows which begin and end in the Seattle Metro Area. Using the areas of the intersections between each block group and parcel, we calculate the share of each block group's area that falls into a particular parcel. We distribute a block group's residents

[39] We visually inspect for the latter condition, generating a list of 30 bilateral pairs which violate it. These pairs are removed from the sample after an earlier filtering, which cuts to the sample to only those bilateral pairs which are between nodes that are in contiguous parcels.

Figure F.1: Route Complexity: An Example



Notes: This figure provides an example of how the instrument for traffic based on the route complexity is constructed for the Seattle road network. On this link, there are five turns and 19 intersections.

and labor force among its intersections according to these shares, and then aggregate those intersections up to the parcel level to find the number of residents and the size of the labor force for each parcel.

To distribute commuting flows, we estimate the commuting flows between block group-parcel intersections by multiplying total commuting flows between two block groups by the shares of the origin block group and the destination block group taken up by the origin intersection and the destination intersection. We then aggregate these commuting flows up to the parcel-to-parcel level.

F.2.5 Instrument Construction

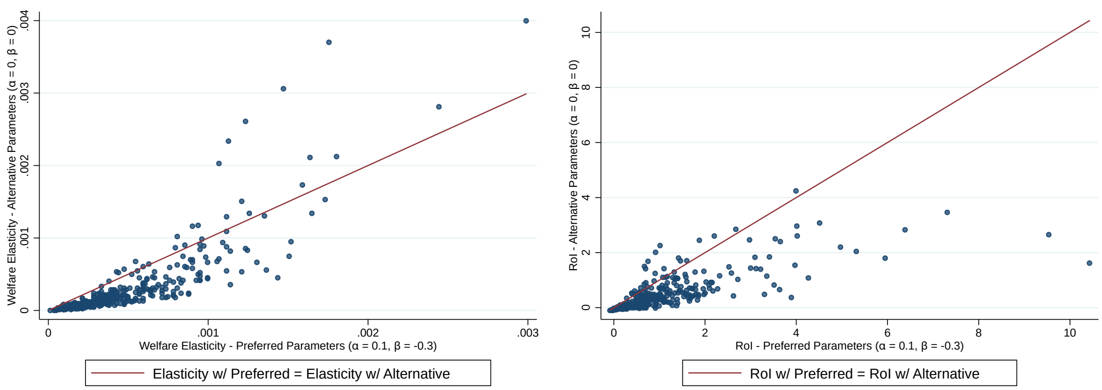
For our IV estimation of the congestion parameter in Seattle, we rely upon the number of turns along a route, conditional on the number of intersections traversed and origin and destination fixed effects, as an instrument. We define any deviation from the current bearing of the route by more than 30 degrees, in either direction, as a turn, and we use the Global Turn Delay within ArcMap's Network Analyst to count the number of turns along the least-cost route between two nodes. We also use the Network Analyst to count the number of intersections traversed. The below figure presents an example of how this process works and what kind of data it results in. Between these two nodes, the least-cost path makes five turns and traverses 19 intersections.

G Online Appendix: Alternative Parameter Constellations

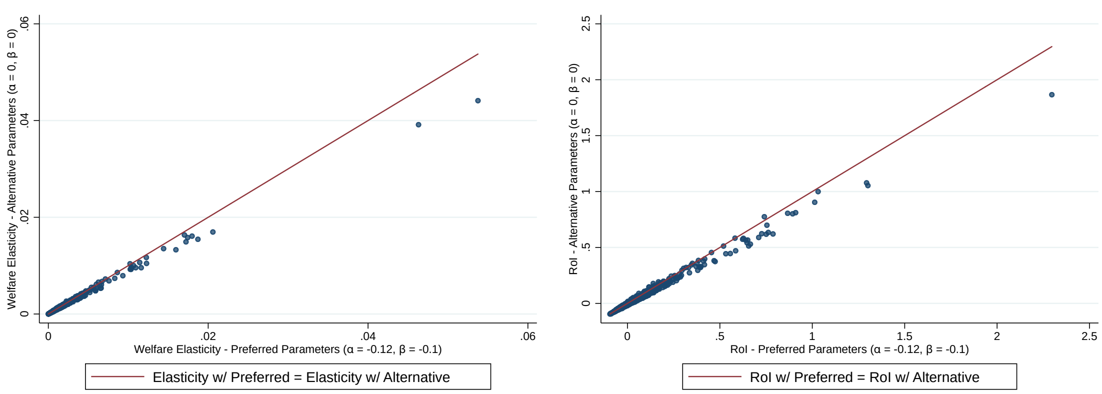
In the section, we compare the estimated welfare elasticities and returns on investment for each segment of the U.S. highway network and the Seattle road network to equivalent results under three different parameter constellations: (1) no externalities ($\alpha = \beta = 0$); (2) lower trade elasticity ($\theta = 4$); and (3) greater traffic congestion, which we calculate by estimating δ_0 from a gravity regression of either trade or commuting on travel times. We summarize the results in three figures, corresponding to each of the three parameter constellations.

Figure G.1: Alternative parameter constellation: No externalities

(a) U.S. Highway Network



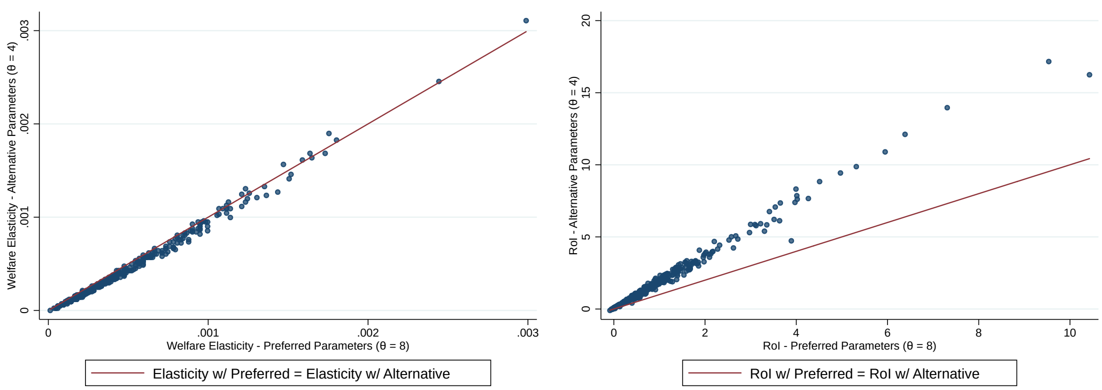
(b) Seattle Road Network



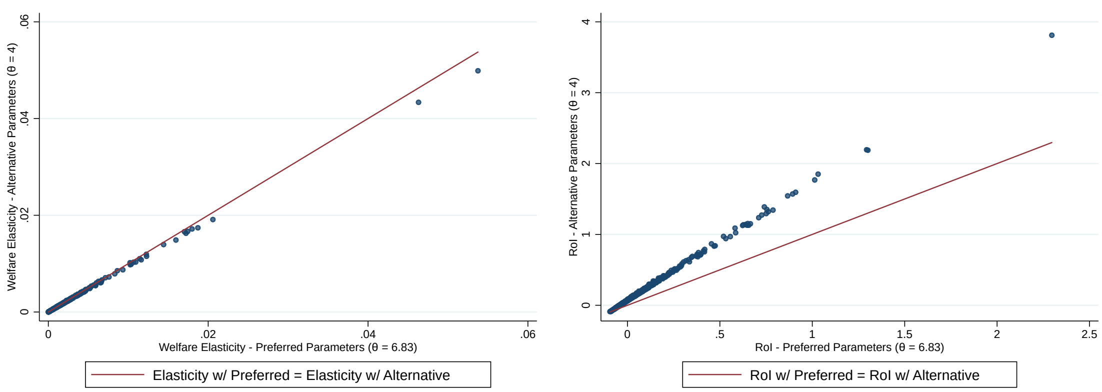
Notes: This figure compares the welfare elasticity (on the left) and return on investment (on the right) elasticity for each link in the U.S. highway network (panel a) and the Seattle road network (panel b) calculated using our preferred parameter constellation (on the x-axis) to an alternative parameter constellation where we assume no externalities, i.e. $\alpha = \beta = 0$, (on the y-axis).

Figure G.2: Alternative parameter constellation: Lower gravity elasticity

(a) U.S. Highway Network



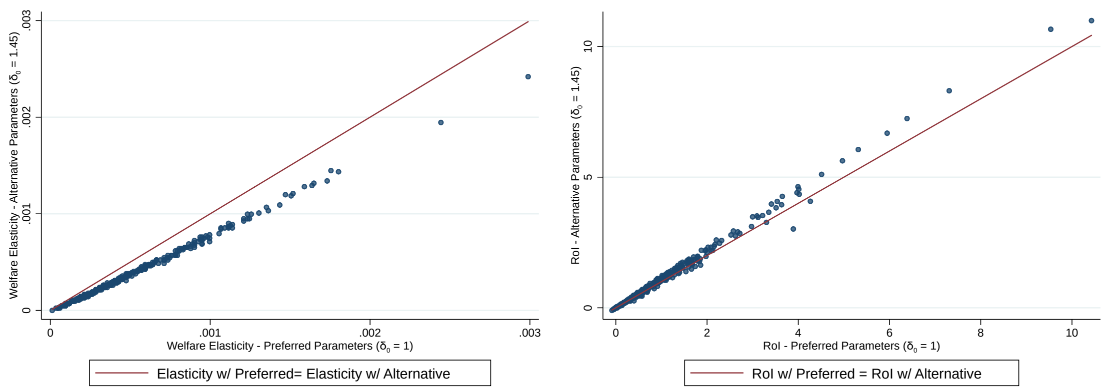
(b) Seattle Road Network



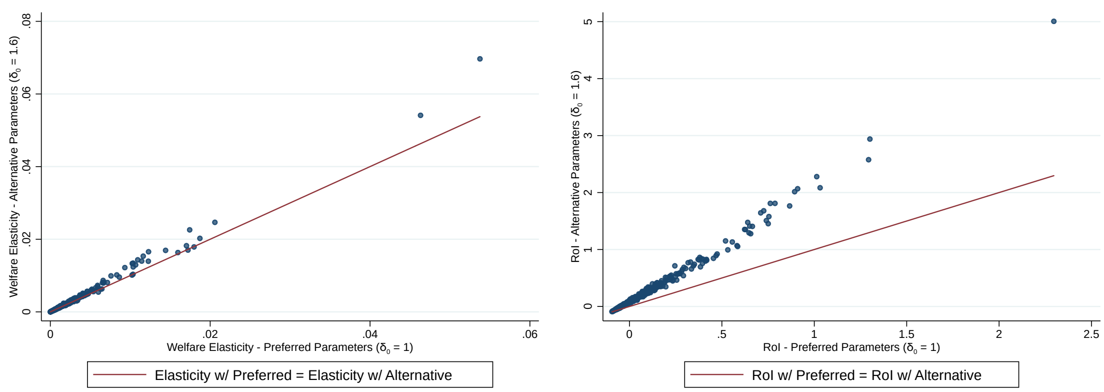
Notes: This figure compares the welfare elasticity (on the left) and return on investment (on the right) elasticity for each link in the U.S. highway network (panel a) and the Seattle road network (panel b) calculated using our preferred parameter constellation (on the x-axis) to an alternative parameter constellation where we assume a smaller gravity elasticity, i.e. $\theta = 4$, (on the y-axis).

Figure G.3: Alternative parameter constellation: Greater traffic congestion

(a) U.S. Highway Network



(b) Seattle Road Network



Notes: This figure compares the welfare elasticity (on the left) and return on investment (on the right) elasticity for each link in the U.S. highway network (panel a) and the Seattle road network (panel b) calculated using our preferred parameter constellation (on the x-axis) to an alternative parameter constellation where we assume greater traffic congestion, i.e. higher λ, (on the y-axis). The λ is calculated here by estimating $\delta_0\theta$ based on a gravity regression of trade (panel a) or commuting (panel b) flows on travel times, rather than setting $\delta_0\theta = 1$, as above.