Review of Economic Studies (2022) **89**, 2911–2957 https://doi.org/10.1093/restud/rdac001

Advance access publication 1 February 2022

The Welfare Effects of Transportation Infrastructure Improvements

TREB ALLEN
Dartmouth and NBER

and

COSTAS ARKOLAKIS
Yale and NBER

First version received January 2019*; Editorial decision October* 2021*; Accepted January* 2022 (*Eds.*)

Each year in the US, hundreds of billions of dollars are spent on transportation infrastructure and billions of hours are lost in traffic. We develop a quantitative general equilibrium spatial framework featuring endogenous transportation costs and traffic congestion 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 US 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.

Key words: traffic, congestion, optimal routing, trade, urban, commuting, transportation.

JEL Codes: F1, R12, R13, R41, R42, H54

1. INTRODUCTION

More than a trillion dollars is spent on transportation infrastructure across the world each year (Lefevre, Leipziger and Raifman, 2014). In the US alone—where annual spending on highways exceeds $150 billion—the average driver spends an average of 42 h 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

The editor in charge of this paper was Adam Szeidl.

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 develop a new tractable spatial framework featuring endogenous transportation costs and 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 US 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 *et al*. (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 US 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 US 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 US 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 US 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 article 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 Ahlfeldt *et al*. (2015), Tsivanidis (2018), Heblich, Redding and Sturm (2020) in an intra-city context, and Monte, Redding and Rossi-Hansberg (2018) combining intra-city and inter-city analyses; 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 modelling. Our analysis retains the key analytical benefits of that previous work but also provides a comprehensive framework to analyse 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 article 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 productivity and amenity externalities and/or traffic congestion given the absence of congestion tolls means the equilibrium is (generically) inefficient. 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]

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)).

The remainder of the article 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 US 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; we refer the interested reader to Supplementary Appendix B for detailed derivations of the results that follow in this section.

For both models, we posit the following geography. Suppose the economy consists of a finite number of locations $i \in \{1,\dots,N\} \equiv \mathcal{N}$ arrayed on a network and inhabited by $\bar{L}$ individuals. Mathematically, this network 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. $r \equiv \{i = r_0, r_1, \dots, r_K = j\}$, 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 labour is mobile, as in Allen and Arkolakis (2014) and Redding (2016).

2.1.1. Setup.

An individual residing in location i supplies her endowed unit of labour 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

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 Supplementary Appendix D.1, we consider an alternative framework where costs incurred travelling through the network are additive and show that one can derive a similar expression for the endogenous transportation costs below.

elasticity of substitution preferences with elasticity of substitution $\sigma \geq 0$. Labour is the only factor used in the production and shipment of goods. Let Y^W and $\bar{L}$ denote the total income and total labour endowment in the economy, respectively; in what follows, we choose average per-capita income as our numeraire, that is, $Y^W/\bar{L}=1$, which implies that the value of trade is measured in average units of labour.

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 is subject to idiosyncratic productivity shocks $\varepsilon_{ij,r}(\nu)$, meant to capture the various uncertainties that production and shipping are subject to. Under perfect competition the price of good ν in destination $j \in \mathcal{N}$ from origin $i \in \mathcal{N}$ along route $r \in \Re_{ij}$ is

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

Individuals in destination j then purchase each good $\nu \in [0,1]$ from the cheapest source (i.e. location-route). 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 $1/A_i$, where A_i captures an origin-specific efficiency, and shape parameter θ, which regulates the (inverse of) shock dispersion.[4]

The main innovation in our setup 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 modelling 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.[5] 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. Avoiding corner solutions by adding such noise is cited as the original impetus for the Eaton and Kortum (2002) framework and the Frechet assumption allows us to further retain the tractability and extend the analytical solutions of that framework in the presence of traffic.[6]

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 Supplementary Appendix D.2). Instead, what is enormously helpful (and which the simultaneous choice over locations and routes ensures) is that agents' elasticities of substitution among locations and among routes are the same. Deviations from this

4. Papers such as Eaton and Kortum (2002), Ramondo and Rodríguez-Clare (2013), and Lind and Ramondo (2018) illustrate how correlated Frechet shocks could be added to a multi-regional analysis. The idea behind such an extension is that some routes may be subject to common underlying conditions and thus subject to correlated shocks. We abstract from such consideration for tractability purposes and because it requires the need of additional (correlation) parameters to be estimated.

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

6. 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 US highways) and 0.98 (for the Seattle road network).

assumption—while computationally straightforward—come so at the loss of substantial analytical tractability and ensuing economic insight.[7]

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

$$A_i = \bar{A}_i L_i^{\alpha},\ u_i = \bar{u}_i L_i^{\beta}, \tag{1}$$

where $\bar{A}_i > 0$ and $\bar{u}_i > 0$ are the *local geography* of productivity and amenities, and $\alpha, \beta \in \mathbb{R}$ govern the strength of the productivity and amenity externalities, respectively. 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. For example, the parameter α can be considered as capturing entry externalities as in Krugman (1991), which lead to more concentration of economic activity, and the parameter β negative amenity spillovers or the presence of a housing market, which lead to dispersion of economic activity. We will contrast the implications of these (now standard) spillovers to the (new) traffic congestion spillovers below.

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. Given the Frechet assumption, the probability that $j \in \mathcal{N}$ purchases good $\nu \in [0,1]$ from $i \in \mathcal{N}$ along route $r \in \mathfrak{R}_{ij}$, $\pi_{ij,r}$, can be written as:

$$\pi_{ij,r} = \frac{(w_i/A_i)^{-\theta} \left(\prod_{l=1}^{K} t_{r_{l-1},r_l}^{-\theta}\right)}{\sum_{k \in \mathcal{N}} (w_k/A_k)^{-\theta} \sum_{r' \in \mathfrak{R}_{kj}} \prod_{l=1}^{K} t_{r'_{l-1},r'_l}^{-\theta}}. \tag{2}$$

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:

$$X_{ij} = \sum_{r \in \mathfrak{R}_{ij}} \pi_{ij,r} E_j = \frac{\tau_{ij}^{-\theta} (w_i/A_i)^{-\theta}}{\sum_{k \in \mathcal{N}} \tau_{kj}^{-\theta} (w_k/A_k)^{-\theta}} E_j, \tag{3}$$

where:

$$\tau_{ij} \equiv \left(\sum_{r \in \mathfrak{R}_{ij}} \left(\prod_{l=1}^{K} t_{r_{l-1},r_l}^{-\theta} \right) \right)^{-\frac{1}{\theta}} \tag{4}$$

is the *transportation costs* from i to j. Note that expression (3) 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.

7. 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). Such deviations may arise if, for example, substitution among production locations is more difficult than substitution among alternative routes (or vice-versa). Such an example and intuition of where these assumptions exactly come to play for our results is provided in Supplementary Appendix D.3 based on the Armington model.

2.1.3. Market access and gravity. While (3) 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, as in Anderson and Van Wincoop (2003) and Redding and Venables (2004). To do so, we first impose two equilibrium market clearing conditions: (1) total income Y_i in each location is equal to its total sales; and (2) total expenditure E_i in each location is equal to its total purchases:

$$Y_i = \sum_{j=1}^{N} X_{ij}, \; E_i = \sum_{j=1}^{N} X_{ji}. \tag{5}$$

We can re-write the gravity equation (3) as follows:

$$X_{ij} = \tau_{ij}^{-\theta} \times \frac{Y_i}{\Pi_i^{-\theta}} \times \frac{E_j}{P_j^{-\theta}}, \tag{6}$$

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

$$\Pi_i \equiv \left(\sum_{j=1}^{N} \tau_{ij}^{-\theta} E_j P_j^{\theta} \right)^{-\frac{1}{\theta}} = A_i L_i Y_i^{-\frac{\theta+1}{\theta}}, \tag{7}$$

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

$$P_j = \left(\sum_{i=1}^{N} \tau_{ij}^{-\theta} Y_i \Pi_i^{\theta} \right)^{-\frac{1}{\theta}}. \tag{8}$$

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:

$$W_j = \frac{w_j}{P_j} u_j, \tag{9}$$

where u_j is an amenity value of living in location $j \in \mathcal{N}$. We assume that there is free labour mobility across locations and we focus in equilibria where welfare equalizes across locations, $W_j = \bar{W}$, and every location is populated.[8]

Combining the definitions in (1), equation (6), the market clearing conditions (5), imposing balanced trade (i.e. $E_i = Y_i$) and welfare equalization (i.e. condition (9)), we obtain the following equilibrium conditions:

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

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

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

where $y_i \equiv Y_i/Y^W$ and $l_i \equiv L_i/\bar{L}$ are the share of total income and total labour 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] Conditional on τ_{ij} this equilibrium system is identical to the one in Allen and Arkolakis (2014) and Redding (2016). In particular, given productivities $\{\bar{A}_i\}$, amenities $\{\bar{u}_i\}$, and transportation costs $\{\tau_{ij}\}$, the $2N$ equations (10) and (11) can be solved for the $2N$ equilibrium shares of income $\{y_i\}$ and labour $\{l_i\}$ in all locations. However, 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 *et al*. (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 the wage in the workplace, w_j; the amenity value of residence, u_i; the time spent commuting; and an idiosyncratic (Frechet distributed with shape parameter θ) route-, origin-, and destination-specific term, $\varepsilon_{ij,r}(\nu)$:

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

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 *et al*. (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 increases 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 labour 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 labour $w_j = A_j$.

9. That χ combines both the equilibrium welfare $\bar{W}$ and the aggregate population $\bar{L}$ demonstrates that 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, that is, 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—that is, 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.

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 (4) 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.

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

$$A_i = \bar{A}_i \left(L_i^F\right)^{\alpha}, \; u_i = \bar{u}_i \left(L_i^R\right)^{\beta}, \tag{14}$$

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.

2.2.3. Market access and gravity. We can now 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:

$$L_i^R \equiv \sum_j L_{ij}, \; L_j^F \equiv \sum_i L_{ij}. \tag{15}$$

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

$$L_{ij} = \tau_{ij}^{-\theta} \times \frac{L_i^R}{\Pi_i^{-\theta}} \times \frac{L_j^F}{P_j^{-\theta}}, \tag{16}$$

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:

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

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:

$$P_j = \left(\sum_j \tau_{ij}^{-\theta} L_i^R \Pi_i^\theta \right)^{-\frac{1}{\theta}} = w_j \left(L_j^F \right)^{-\frac{1}{\theta}} \left(\frac{\bar{L}}{\bar{W}^\theta} \right)^{\frac{1}{2\theta}}. \tag{18}$$

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. Substituting equations (14) into the commuting gravity equation (13) and imposing the equilibrium market clearing conditions (15) 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 $\{l_i^R\}$ and where they work $\{l_i^F\}$. 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 aggregate outcomes of 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 (4)); (2) identical gravity expressions for the bilateral flow of goods/commuters as a function of bilateral costs and market access (summarized in equations (6) and (13), respectively); and (3) 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 we develop, which we turn to next.

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).

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. We refer the interested reader to Appendix A for detailed derivations of the results that follow in this section.

3.1. *Transportation costs*

Both the economic geography and urban models yield transportation costs of the form given in equation (4). By explicitly enumerating all possible routes, equation (4) can be written in matrix notation as follows:[11]

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

where $\mathbf{A}\equiv\left[t_{ij}^{-\theta}\right]$, that is, $\mathbf{A}$ is an $N\times N$ matrix with (i,j) element $t_{ij}^{-\theta}$ (not to be confused with the vector of productivities) and $\mathbf{A}^K=\left[A_{ij}^K\right]$, that is, 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 $\mathbf{B}=\left[b_{ij}\right]$ 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 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 $\left\{\tau_{ij}\right\}_{i,j\in\mathcal{N}^2}$, 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]

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).

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

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.[15]

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

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

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 travelling through link (k,l) along the way from i to j relative to the unconstrained cost of travelling 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 value of goods shipped (in the economic geography model) or the total number of commuters (in the urban model) over the link (k,l). 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 ((6) in the economic geography model or (16) in the urban model) with the link intensity equation (23), we obtain the following expression for equilibrium traffic flows:

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

complexity of $O(N^2 \log N)$. 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 s using Djikstra's algorithm—a three orders of magnitude improvement.

15. See Appendix A.2 for a detailed derivation.

Equation (24) offers a *gravity equation for traffic*, where all determinants of the flow of traffic along link (k,l) are fully summarized by the cost to travel along the link (t_{kl}) and the economic conditions at the beginning and end of the link. It shows the 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. 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 travelling 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 modelling of traffic flows, we now suppose that the direct cost of travelling 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 travelling 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 $\bar{\mathbf{T}} \equiv \left[\bar{t}_{kl}\right]$ 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 (1) and (14), 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 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.

It is important to note that the measure of traffic—and hence traffic congestion—is in the same units that we measure bilateral flows, that is, in the economic geography model, traffic is measured in the value of goods flowing over a link, whereas in the urban model, traffic is measured in the quantity of commuters flowing over a link. There are several advantages to this approach. First, by measuring traffic in the same units that we measure bilateral flows, we generate a close connection between the (new) gravity equation (24) for traffic on a link and the (traditional) gravity equation for flows between an origin and destination (i.e. equations (6) and (16) for the economic geography and urban models, respectively). Second, as we will see below, this close connection allows us to derive analytical equilibrium conditions for the distribution of

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.

economic activity solely as a function of the model fundamentals by solving the same number of equations for the same number of unknowns despite the additional complicated feed-back loop that the presence of traffic congestion generates. Third, retaining the same units for traffic and bilateral flows—along with the assumed log-linear congestion relationship in equation (25)—ensures that the transportation costs between origin and destination remain ad-valorem in the presence of traffic congestion, that is, our framework follows the large literature focusing on so-called iceberg transportation costs.

These advantages notwithstanding, however, a reasonable objection that applies to the economic geography framework is that traffic congestion actually is increasing in the quantity rather than the value of trade: for example, a truck carrying cheap apples generates the same traffic congestion as one carrying expensive apples. In the Supplementary Appendix D.4, we show how the economic geography framework can be easily altered to to assume instead that traffic (and traffic congestion) are measured in the quantity of labour used to produce the goods and Supplementary Appendix D.5 for the case where traffic is measured in the quantity of goods. In both cases, we show that equilibrium traffic flows also follow a gravity equation nearly identical to that of equation (24), differing only in that the market access measures are quantity-based rather than value-based. However, the need to simultaneously consider both quantity- and value-based market access measures increases the complexity of the equilibrium system, for example, increasing the set of endogenous variables (and systems to solve) from $2N$ to $3N$ when traffic congestion depends on the quantity of labour used to produce the goods.

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.

4.1. *General equilibrium with traffic*

We begin by formally defining our equilibrium: given a local geography $\{\bar{A}_i, \bar{u}_i\}_{i \in \mathcal{N}}$, an aggregate labour 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 $\{l_i^F, l_i^R\}_{i\in\mathcal{N}}$ 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 economic activity ensures markets clear, that is, 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, that is, 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, that is, 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, that is, $y_i > 0$ 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.

Before deriving the new equilibrium system, two remarks are in order. First, in the absence of traffic congestion—that is, $\lambda=0$—then conditional on the equilibrium transportation costs $\{\tau_{ij}\}$ that arise from agents optimal routing decision, the equilibrium is equivalent to the standard spatial setup upon which it is based, that is, our framework tractably nests the standard no-congestion case. Second, with traffic congestion—that is, $\lambda > 0$—the equilibrium will differ from the no-congestion case, as the level of economic activity across space determines the cost of shipping in each link through traffic congestion, differences which we discuss further below. This also implies that the counterfactual predictions of our new setup with traffic congestion cannot be determined by substituting unobserved transportation costs with observed data following the "exact hat" approach of Dekle *et al.* (2008), as now τ_{ij} is endogenous and depends on the entire network of connections through traffic and not just on bilateral flows. We nevertheless devise a new procedure, same in spirit to their exercise, but which instead replaces the need of knowledge of the entire network of connections with the use of traffic data. We discuss this in Sections 5.1 and 5.2.

Consider first the economic geography model. Recall that equations (10) and (11) characterize the equilibrium distribution of population and income as a function of the endogenous transportation costs $\{\tau_{ij}\}$, that is, 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 and then substitute the endogenous transport costs using equations (24),(26), (27), yielding:[17]

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

17. See Appendix A.3 for a detailed derivation of the equilibrium system for both the economic geography and urban models.

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

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 dimensionality of resulting equilibrium system is not larger than the typical system treating transportation costs as exogenous, as the number of equations and number of unknowns remains the same. That allows us to make some progress in characterizing their positive properties (existence and uniqueness), which we turn to next.

4.2. *Existence and uniqueness of equilibrium with traffic*

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 *et al.* (2020) can be extended to analyse the properties of such an equilibrium.[18] We 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 now provide conditions regarding existence and uniqueness in the following proposition.

18. In the absence of traffic congestion, the equilibrium (e.g. equations (10) and (11) in the economic geography model) is an example of a system of non-linear integral equations known as a Hammerstein equation of the second kind, see for example, Polyanin and Manzhirov (2008). Since the results hold for Lebesgue integrals they also apply for a discrete set of locations as discussed in Allen and Arkolakis (2014). Such systems, however, do not admit the inclusion of an endogenous additive term, as in (28) and (29) and also in (30) and (31).

Proposition 1. *For any strictly positive local geography* $\left\{\bar{A}_i > 0, \bar{u}_i > 0\right\}_{i\in\mathcal{N}}$, *aggregate labour endowment* $\bar{L} > 0$, *strongly connected infrastructure network* $\bar{\mathbf{T}} \equiv \left[\bar{t}_{kl}\right]$, *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 Supplementary Appendix C.1. □

Part 1 of Proposition 1 relies on showing that the equilibrium system defined by Equations (28) and (29) for the economic geography model and equations (30) and (31) for the urban model can be transformed into a continuous operator on a compact space so that Brouwer's fixed point theorem applies; whereas Part 2 uses a bounding argument in the spirit of Karlin and Nirenberg (1967) and Allen *et al.* (2020) to show that a (different) transformation of the respective systems would generate a contradiction under the reported parameter constellations.

Despite the added complexity of endogenous traffic congestion (and the involved nature of the proofs), the sufficient conditions for uniqueness in the economic geography model provided in part (a) of the Proposition are identical to those of an economic geography model with exogenous transportation costs, provided by Allen and Arkolakis (2014): the sum of the productivity and amenity externalities must be (weakly) negative to ensure a unique equilibrium, that is, on net the forces that cause dispersion need to dominate the forces that cause concentration. In the urban model, we achieve a similar result but since we do not impose symmetry the productivity and amenity spillovers must satisfy a related condition individually, rather than combined).[19] 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 *lower the values* of the productivity and amenity externalities must be to satisfy these sufficient conditions for 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.

4.3. *Traffic congestion and 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 labour endowment $\bar{L}$, that is, both (standard) spatial models

19. We should note that for the Allen and Arkolakis (2014) model these sufficient conditions are also necessary, assuming arbitrary geography of trade costs (see Theorem 1(iii) in Allen, Arkolakis and Li (2020)). Unfortunately, such a characterization of necessary conditions is not possible in the presence of traffic congestion. It thus remain a possibility that weaker sufficient conditions can be proven where the traffic congestion parameter, λ, plays a role in such characterization.

are *scale invariant*.[20] In the presence of traffic congestion, however, the equilibrium distribution of economic activity does depend on the size of the aggregate labour endowment $\bar{L}$, that is, the equilibrium is *scale dependent*. As is evident from equations (28) and (29) (in the economic geography model) and equations (30) and (31) (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 labour 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 $c > 1$, that is, suppose t_{kl} increases to ct_{kl}, changes equilibrium transportation costs (holding constant traffic congestion fixed). Differentiating equation (21) around $c = 1$ yields:[21]

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

that is, 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. Figure 1(a) depicts the equilibrium distribution of economic activity in the absence of congestion forces (i.e. $\lambda = 0$). Locations in the centre of the grid with better market access enjoy greater equilibrium economic activity (as indicated by taller "buildings"), and links in the centre of the grid experience greater traffic (as indicated by their colour), as they are more heavily used to travel through the network.

In Figure 1(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 travelled 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 centre of the city and rises in the outskirts: that is, traffic congestion forces agents out of the centre of the city and into the suburbs.

In Figure 1(c) and (d), we increase the size of the economy from $\bar{L} = 100$ to $\bar{L} = 1000$ (Figure 1(c)) and $\bar{L} = 10,000$ (Figure 1(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 centre of city is the worse affected: the resulting gridlock induces a reallocation of economic activity away from the centre and toward the edges, further amplifying the move to the suburbs.

20. 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.

21. See Supplementary Appendix B.3 for a detailed derivation.

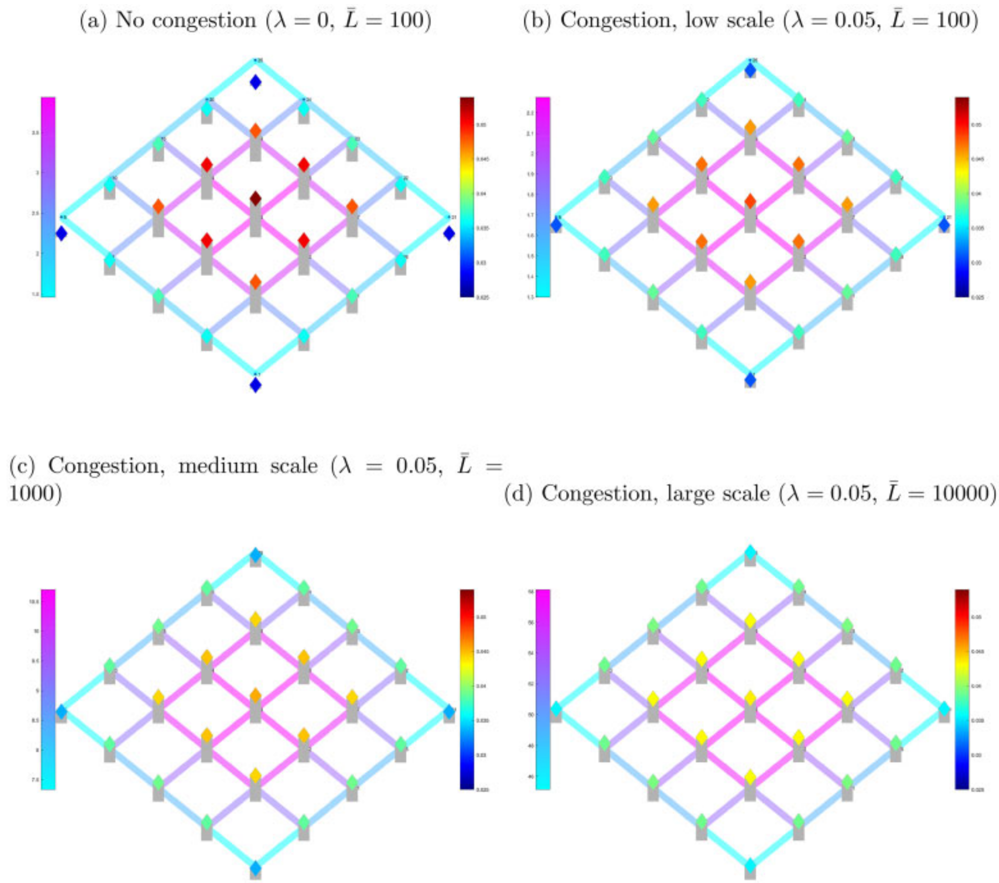


FIGURE 1

Traffic congestion and the distribution of economic activity

Notes: This figure shows how traffic congestion (λ) and the scale of the economy $(\bar{L})$ shapes the distribution of economic activity within an example 5×5 grid network using the urban model. The height of the buildings (and the rooftop colours, associated with the colour bar on the right) indicate the equilibrium residential population (L_i^R) at each location in the city, and the colour of each link (associated with the colour 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.

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 6 and 16, 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:[22]

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

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

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 $\mathbf{C}^L \equiv \left(\mathbf{D}^L - \Xi\right)^{-1}$, $\mathbf{D}^L$ 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 $\Xi \equiv \left[\Xi_{ij}\right]$.

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, Ξ.[23] 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 discussed below), it provides an out-of-sample test of the model predictions about traffic flows. In addition, if trade/commuting data are not available, but traffic data is (e.g. as in 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*

To evaluate the welfare impact of transportation infrastructure improvements in the presence of traffic congestion, we next analyse how to conduct counterfactuals. To do so, we follow the "exact hat algebra" approach pioneered by Dekle *et al.* (2008), where we denote with hats the change in variables, $\hat{y}_i \equiv \frac{y_i'}{y_i}$, where we denote with prime the counterfactual outcome. We summarize the result in the following proposition.

Proposition 2. *Suppose an observed economy has infrastructure network* $\bar{\mathbf{T}} \equiv \left[\bar{t}_{kl}\right]$ *and is in equilibrium. Consider any change in the underlying infrastructure network denoted by* $\hat{\bar{t}}_{kl}$. *Given observed traffic flows,* $\left[\Xi_{ij}\right]$, *economic activity in the geography* $\left(Y_i, E_j\right)$ *or urban model* $\left(L_i^R, L_j^F\right)$

22. See Supplementary Appendix B.4 for a detailed derivation.

23. By imposing symmetry, one can also recover the transportation costs of traversing each link of the infrastructure network as summarized by the adjacency matrix $\mathbf{A} \equiv \left[t_{kl}^{-\theta}\right]$; this procedure—which applies a variant of the methodology developed by Head and Ries (2001) to recover bilateral trade costs from bilateral trade flows to equations (34) and (35). In combination with equation (23), this then allows us to calculate the link intensity π_{ij}^{kl} using only observed traffic and economic activity data, something we return to below. We use this procedure—discussed in detail in Supplementary Appendix B.4—in creating Figure 4 below.

and parameters $\{\alpha,\beta,\theta,\lambda\}$*, the equilibrium change in economic outcomes* $\left(\hat{y}_i,\hat{l}_i,\hat{\chi}\right)$ *is the solution the following system of equations:*

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

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

for the economic geography model and as:

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

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

for the urban model. Moreover, existence and uniqueness of the counterfactuals are given by the same conditions as in Proposition (1).

Proof. See Supplementary Appendix C.2. □

Proposition 2 says that given observed traffic flows and the observed distribution of economic activity—and knowledge of the model parameters $\{\theta,\alpha,\beta,\lambda\}$—it is possible 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.[24] Note that equations (36)–(39) all say that some log linear combination of endogenous changes in location i depend on a weighted average of a (different) log linear combination of endogenous changes in location i and a (third) log linear combination of endogenous changes in all location j, where the weights are determined by the relative size of observed local economic activity and traffic flows. Loosely speaking, this locations with large amounts of traffic flows to i will play a greater role in determining the counterfactual outcomes in i, all the more so if these traffic flows are large relative to the economic activity in i. It is worth emphasizing that conducting counterfactuals using this result requires easily observed traffic flows along links in the network, instead of potentially harder to observe bilateral trade or commuting flows between origins and destinations upon which

24. Supplementary Appendix E describes the algorithm used to solve equations (19) and (20) given these ingredients.

traditional implementations of the Dekle *et al.* (2008) "exact hat" algorithm rely (see e.g. Redding, 2016; Caliendo, Parro, Rossi-Hansberg and Sarte, 2018; Adao, Arkolakis and Esposito, 2021).

The second part of Proposition 2 says that existence and the sufficient conditions for uniqueness for the counterfactuals are the same as for the system in levels. This result arises from the fact that the systems of equations that determine the counterfactual outcomes in changes are mathematically equivalent to their level variants above, where the local geography and infrastructure matrix are simply replaced with shares that depend only on observed traffic flows and the observed distribution of economic activity.

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.[25] As a result, we can write t_{kl} as a function of the distance of the link and the speed of travel on the link:

$$t_{kl} = \left(\text{distance}_{kl} \times \text{speed}_{kl}^{-1}\right)^{\delta_0}, \tag{40}$$

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 for example, Disdier and Head (2008) and Chaney (2018).[26]

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

$$\text{speed}_{kl}^{-1} = m_0 \times \left(\frac{\Xi_{kl}}{\text{lanes}_{kl}}\right)^{\delta_1} \times \varepsilon_{kl} \tag{41}$$

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.[27] First, combined with equations (40), and (41) immediately implies:

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

25. 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 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).

26. In Supplementary 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.

27. 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, that is, while we follow Vickrey (1967) in considering a log-linear approximation of the impact of congestion on travel time, our approximations centres around an inverse speed of zero rather than the free-flow rate of travel.

where $\bar{t}_{kl} \equiv \text{lanes}_{kl}^{-\delta_0\delta_1} \times (\text{distance}_{kl} \times m_0 \times \varepsilon_{kl})^{\delta_0}$ 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:

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

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 $\frac{\partial \ln \Xi_{kl}}{\partial \ln \text{lanes}_{kl}} = \frac{\lambda\theta}{1+\lambda\theta}$, 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).[28]

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:

$$\ln \text{speed}_{kl}^{-1} = \ln m_0 + \delta_1 \ln\left(\frac{\Xi_{kl}}{\text{lanes}_{kl}}\right) + \ln \varepsilon_{kl}, \tag{43}$$

that is, 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.[29] 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 US 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.

28. It is important to note that this is the partial elasticity of traffic to additional lanes, whereas Duranton and Turner (2011) empirically evaluate the total elasticity of traffic to additional lanes, including the resulting general equilibrium changes in the spatial distribution of economic activity.

29. Alternatively, we could have calibrated δ_1 by adapting the estimates of Couture, Duranton and Turner (2018). The authors estimate a very similar relationship using instrumental variables, but estimate a different elasticity for traffic and number of lanes. Their preferred estimates imply that the elasticity on traffic is larger than the elasticity on lanes, but the difference is typically very small.

6.1. *Traffic across the country: the US highway network*

The US National Highway System is the 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 projects 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 US population. Such knowledge is crucial for appropriately targeting future infrastructure investments.

Our strategy to estimate the welfare impact of improvements to the US Highway System is straightforward: for each segments of the network, we will use equations (36) and (37) from Proposition 2 for the economic geography variant of our approach to estimate the aggregate welfare impact ($\hat{\hat{W}} = \hat{\chi}^{-\frac{1}{\theta}}$) 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 Supplementary 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 US highway system.[30]

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 results in 228 locations and 704 links between adjacent nodes, where for each link we construct a length-weighted average of AADT and number of lanes. Figure 3(a) 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 US 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.[31] To determine

30. The traffic data are 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.

31. 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 (\$1.923 m), rural-rolling (\$2.085 m), rural-mountainous (\$6.492 m), small-urban (\$3.061 m), small-urbanized (\$3.345 m), large-urbanized (\$5.598 m), and

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 vs. 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 backed out bilateral trade flows using equation (34) and the observed traffic flows. 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 labour. 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. Figure 2(a) 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 for example, 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.[32]

Finally, Figure 4(a) provides an example of intensity of usage of different links for a specific origin and destination pair, Los Angeles, California to New York, New York. As expected, the links that are on very direct routes, such as for example segments of the I-95 interstate near New York, are very intensively used to serve that pair, whereas more indirect links such as highway segments in California north of Los Angeles, have negligible usage.

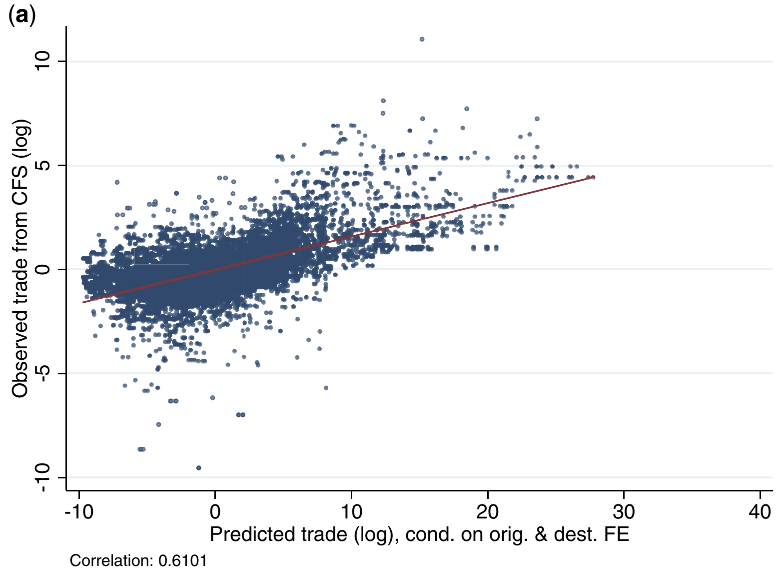
6.1.3. Estimation. We now discuss our choice of the four model parameters $\{\theta, \alpha, \beta, \lambda\}$. As the first three model parameters—the trade elasticity θ, productivity externality α, and amenity externality β—are standard in the economic geography literature, we choose central values from the literature. We set $\theta = 8$ to match previous estimates of the trade elasticity.[33] 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 for example, Allen and Arkolakis (2014).[34] From Proposition 1, this choice of parameter values guarantees the existence of a unique equilibrium.

major-urbanized (\$11.197 m). We are grateful to the experts at the US Department of Transportation Volpe National Transportation Systems Center for their substantial assistance in developing these cost estimates.

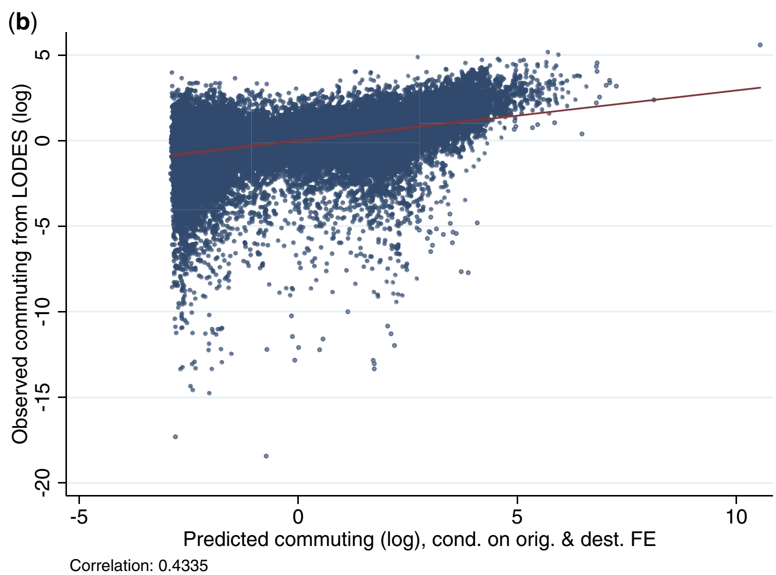
32. The model is able to fit the rapid decline of trade with distance (Hillberry and Hummels, 2008) as Supplementary Appendix Figure F.2 illustrates.

33. Donaldson and Hornbeck (2016) estimate a trade elasticity of 8.22 for US 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.

34. 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 Supplementary 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.



Trade flows in an economic geography model

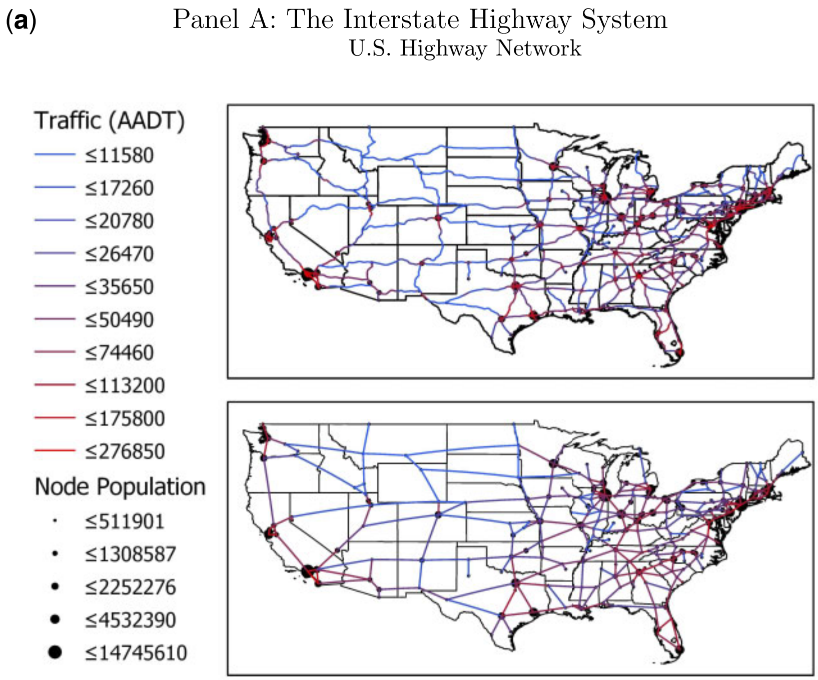


Commuting flows in an urban model

FIGURE 2

Predicting flows using traffic

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.



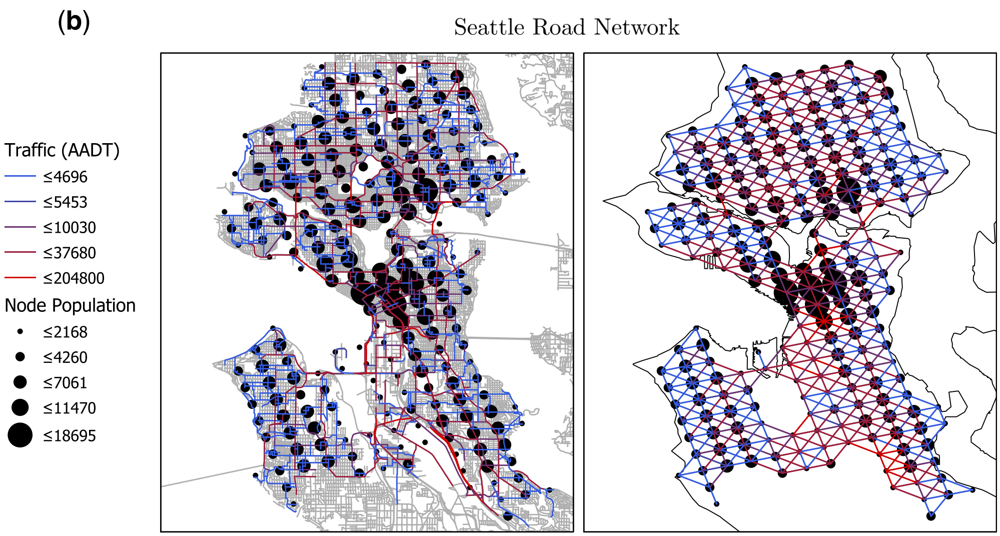
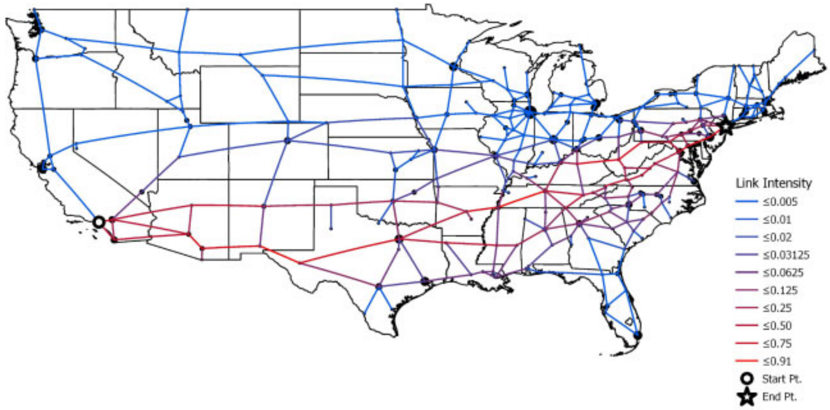


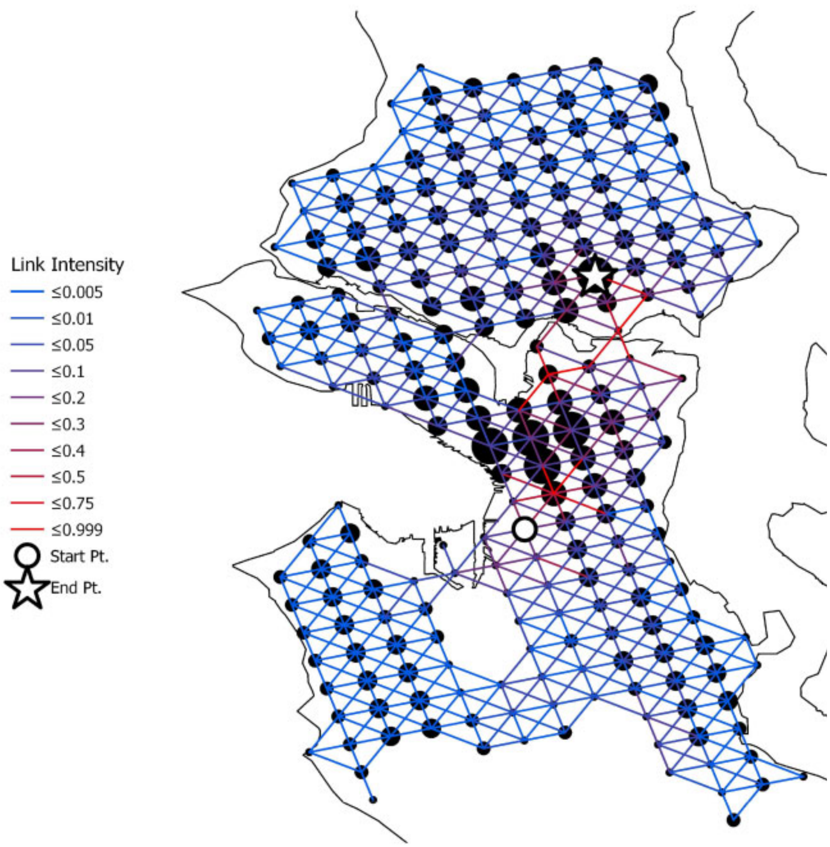
FIGURE 3

Transportation systems and their network representations

Notes: This figure presents the observed transportation network (on the top) and the constructed infrastructure matrix (on the bottom) for the US 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 colour of each link reflects the amount of traffic with red (blue) indicating high (low) levels of traffic. The grey roads in panel (b) are roads not on the least cost route between grid centres.



(a) Economic geography: Los Angeles, CA to New York, NY



(b) Urban model: Safeco Field to the University of Washington

FIGURE 4

Example link intensities

Notes: This figure shows an example of the link intensity π_{ij}^{kl} – i.e. the expected number of traverses of each link in the network – across all chosen routes from Los Angeles, California to New York, New York in panel (a) and from Safeco Field to the University of Washington in panel (b). These link intensities are calculated using only observed data on traffic flows and the economic activity in each location (i.e. no assumptions on model parameters are necessary); see Supplementary Appendix B.4 for details.

TABLE 1
Estimating the strength of traffic congestion

	Panel A: Interstate highway system			
	(1)	(2)	(3)	(4)
	OLS	OLS	IV: 1st stage	IV: 2nd stage
Log congestion	0.106***	0.050***		0.739***
	(0.010)	(0.011)		(0.181)
Log distance			−0.156***	
			(0.033)	
Start-location FE	No	Yes	Yes	Yes
End-location FE	No	Yes	Yes	Yes
F-statistic	116.371	19.987	22.747	16.656
Observations (excl. singletons)	630	630	630	630
Observations (incl. singletons)	704	704	704	704

	Panel B: Seattle road network				
	(1)	(2)	(3)	(4)	(5)
	OLS	IV: 1st stage	IV	IV: 1st stage	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	6.195	5.336	3.084
Observations	1,338	1,338	1,338	1,338	1,338

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: $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

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 US 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 US 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).

Table 1 Panel (a) 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 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 $\delta_1 = 0.739$ (with standard error of 0.181). Recall from above that we set $\delta_0 = 1/\theta$ to match the unit distance elasticity, so this implies $\lambda = \delta_1 \delta_0 = 0.092$, that is, 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.[35]

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 US Highway System using equations (36) and (37) of Proposition 2, i.e. $\frac{1}{2}\left(\frac{\partial \ln \bar{W}}{\partial \ln \bar{t}_{kl}} + \frac{\partial \ln \bar{W}}{\partial \ln \bar{t}_{lk}}\right)$. Figure 5(a) 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? Figure 6(a) compares the welfare elasticity for each segment with and without congestion. From the scatter plot on the right, it is clear that 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. From the map on the left, we see that ignoring traffic congestion overstates the welfare gains from infrastructure improvements the most along highly trafficked segments of the highway system such as I-5 between Los Angeles and San Diego, CA as well as along highway segments around important hubs for intrastate shipping such as those surrounding Atlanta, GA, highlighting the fact 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 a 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 US 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 multiply 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.[36]

35. Our estimate of δ_1 implies a (partial) elasticity of traffic on a segment to additional lane-miles of $\frac{\partial \ln \Xi_{kl}}{\partial \ln lanes_{kl}} = \frac{0.739}{1+0.739} \approx 0.4$, a substantial effect, albeit below the value of one implied by the "fundamental law of road congestion" of Duranton and Turner (2011).

36. 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.

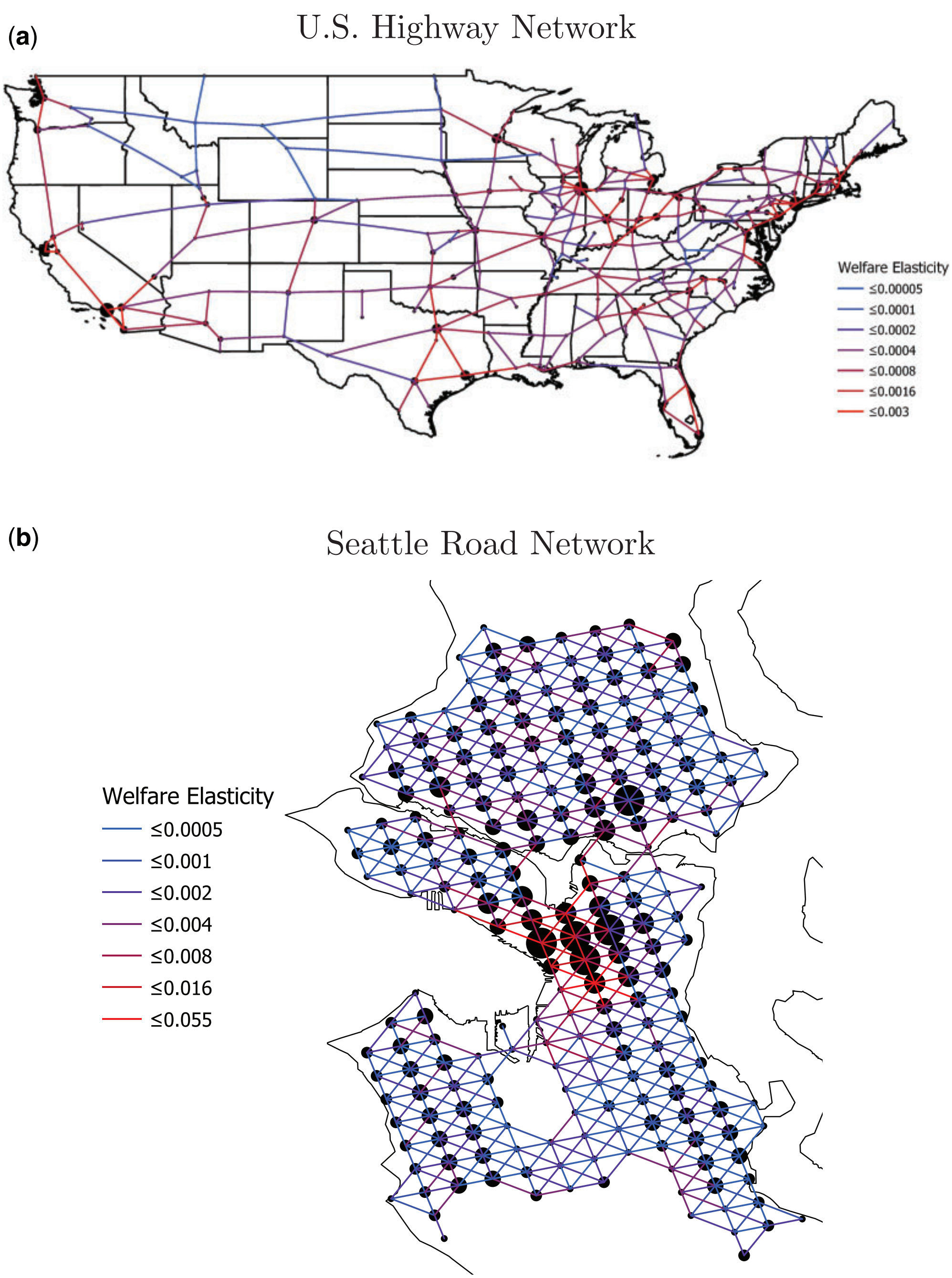


FIGURE 5

Welfare elasticities of infrastructure improvement

Notes: This figure presents the elasticity of aggregate welfare to improving each link in the US Highway Network (panel A) and the Seattle road network (panel B). The colour 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.

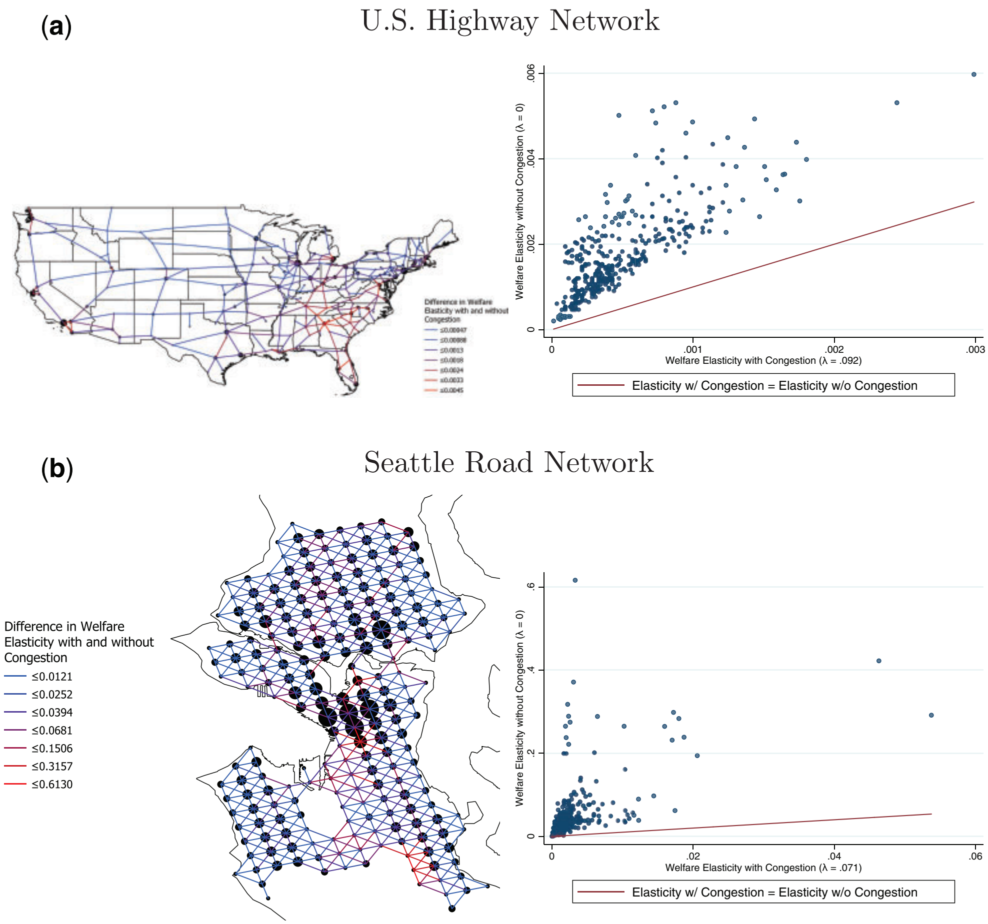


FIGURE 6

Comparing welfare elasticities with and without congestion

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 US highway network (panel a) and the Seattle road network (panel b). The left figure in the panel shows the difference in welfare with and without congestion across links, whereas the right figure panel shows a scatter plot of the two estimated elasticities.

Figure 7(a) reports the annual return on investment (RoI) for each segment of the US 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. Table 2(a) 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 analyse the urban variant of our framework to examine the welfare impacts of transportation infrastructure improvement in Seattle, WA. Seattle provides an ideal test-case

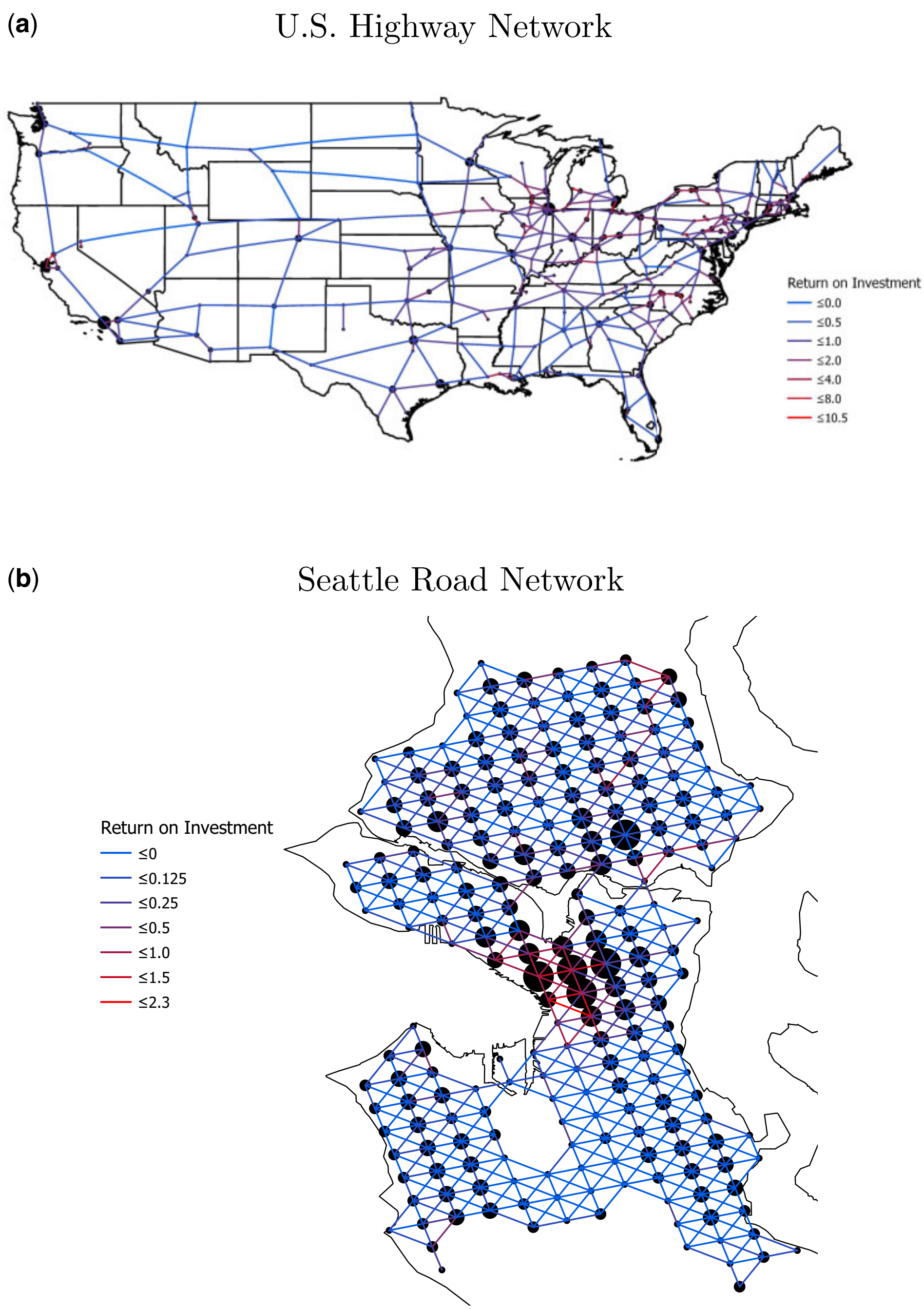


FIGURE 7

Returns on investment of infrastructure improvement

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 colour 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 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 neighbourhood	Home address	Work neighbourhood	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	0.91	3.474	3.44
7	Downtown Seattle	2020 4th Ave	Lower Queen Anne	201 Queen Anne Ave N	0.89	7.973	8.024
8	Capitol Hill	1698 Yale Ave	South Lake Union	1214 E Roy St	0.87	6.596	6.823
9	SoDo	562 1st Ave S	Downtown Seattle	83 Spring St	0.79	3.021	3.404
10	Downtown Seattle	83 Spring St	Downtown Seattle	2020 4th Ave	0.76	5.619	6.518
11	South Lake Union	401 Aurora Ave N	South Lake Union	1214 E Roy St	0.75	4.738	5.545
12	North Seattle	2517 NE 120th St	Cedar Park	13504 WA-522	0.75	2.026	2.381
13	Northeast Seattle	4200 Mary Gates Memorial Dr NE N207	Northeast Seattle	4700 45th Ave NE	0.74	1.261	1.5
14	Downtown Seattle	1001 Columbia St	Downtown Seattle	2020 4th Ave	0.73	5.001	6.052
15	Downtown Seattle	83 Spring St	Capitol Hill	1698 Yale Ave	0.71	4.261	5.259
16	Greater Duwamish	1150 S Atlantic St	Downtown Seattle	805 S Jackson St	0.67	9.416	12.306
17	SoDo	1705 4th Ave S	SoDo	562 1st Ave S	0.66	2.696	3.567
18	Downtown Seattle	1001 Columbia St	Capitol Hill	1698 Yale Ave	0.65	3.802	5.066
19	Cedar Park	13504 WA-522	Olympic Hills	14322 19th Ave NE	0.65	2.373	3.177
20	Downtown Seattle	2020 4th Ave	South Lake Union	1214 E Roy St	0.64	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. As benefits are annual benefits (measured in millions of chained 2012 USD in Panel A and millions of 2016 USD in Panel B), and costs are incurred over a ten year period, the RoIs are annualized returns (i.e. $RoI = (\text{benefit} - 0.1 \times \text{cost})/\text{cost}$, where e.g. 1.2 means an annual 120% return).

for our framework for several reasons, notably: (1) it has some of the worst traffic in the US; (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.[37]

Our strategy for estimating the welfare impacts of improvements to the Seattle road network proceeds analogously to the US 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.

6.2.1. Data. We briefly summarize the data used here; see Supplementary 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.[38] To construct our adjacency matrix of Seattle, we divide Seattle into ~1 sq. mi. grids, place the centre point of each of these grids as a node into ArcGIS Network Analyst, and find the least-cost path between each of these nodes.[39] This gives us a total of 217 nodes, with 1,384 links between adjacent nodes, 1,338 for which we observe traffic.[40] Figure 3(b) 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 labour 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,

37. 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).

38. 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 modelled as flowing from an agents' residence to her workplace—to the (presumed) empirical reality that the agent returns home after work.

39. 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.

40. 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.

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 it, as estimated by the FHA's HERS from Federal Highway Administration (2015).[41] Fifth, for the construction of our instrument, we calculate the number of intersections and turns along each link of the network using the ArcGIS network analyst.

6.2.2. Predicted vs. 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 backed out from equation (35) using 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 Figure 2(b) 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.[42]

Finally, Figure 4(b) shows the intensity of usage of different links for an example commute from Safeco Field to the University of Washington, both on opposite sides of the city centre. As with the interstate highway system, links along the most direct routes are most intensively used. The figure also highlights the fact that different links through the city centre are quite substitutable with each other, with no one link being used more than about half the time, whereas the natural choke-points—e.g. the bridges over Lake Union—are traversed essentially on all routes. In contrast, routes not along the direct route are used negligibly.

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 *et al*. (2015), with $\theta = 6.83$, $\alpha = -0.12$, and $\beta = -0.1$.[43] 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, appropriately instrumented by a demand shifter uncorrelated with the free-flow rate of speed over the link. Unfortunately, the instrument used for the US highway system—distance—is inappropriate in a city setting. There exists enormous variation in the types of roads and speed of travel within

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

42. Traffic predicted commuting flows also predict well the rapid decline of commuting flows with distance; see Supplementary Figure F.2 (b).

43. Ahlfeldt *et al*. (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 Supplementary 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 US highway network, both the patterns of welfare elasticities and the returns on investment are both qualitatively and quantitatively similar to the results presented here.

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*.[44] 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.

Table 1 Panel (b) 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 non-parametrically control for the share of the route that travels over arterial and local roads.[45] Such a procedure compares 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 $\lambda = \delta_1 \delta_0 = 0.071$, that is, 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 US 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.

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 $\frac{1}{2}\left(\frac{\partial \ln \bar{W}}{\partial \ln \bar{t}_{kl}} + \frac{\partial \ln \bar{W}}{\partial \ln \bar{t}_{lk}}\right)$. Figure 5(b) presents our findings. While a reduction in transportation costs on all links are welfare improving, the largest welfare elasticities are greatest in the centre of the city (downtown). Welfare elasticities are also higher for the various choke-points in the road network (oftentimes corresponding to bridges over water).

Figure 5(b) compares these estimated welfare elasticities to those estimated without traffic congestion. As with the US highway system, ignoring congestion would not just result in overestimates of the welfare elasticities, it would also substantially change which links one would identify as having the largest welfare effects. The left figure shows the variation across links in the degree to which one would overestimate welfare gains by ignoring congestion. As is evident, heavily trafficked links near the city centre and along interstate I-5 whose gains fall the most in the presence of traffic congestion. For example, ignoring traffic congestion would cause one to identify a stretch along interstate I-5 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.

44. See Supplementary Appendix Figure F.1 for an example of how the instrument is constructed.

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

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 US 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 vlue to the aggregate welfare gains.[46] We find that improving the average link in Seattle yields an annual return of 16.8% for the residents of the city.[47] Like with the US. highway system, however, there is substantial heterogeneity, with returns varying from less than 25% to more than 250%. Figure 7(b) shows the RoI for each segment; the highest returns are concentrated in the centre of the city. Table 2 Panel b 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 neighbourhood 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.

7. CONCLUSION

This article 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 spatial distribution of economic activity. 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 US highway network and the Seattle road network, congestion matters for, where you improve the road network.

The goal of this paper has been to provide a tractable framework that bridges the gap between the quantitative spatial and transportation economics literatures. Qe 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 the spatial distribution of economic activity? We look forward to fruitful future research on these topics.

Acknowledgments. 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 benefitted from discussions

46. 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.

47. It is interesting to note that the average return on investment is substantially lower in Seattle than in the US highway network. While the different underlying models make a direct comparison difficult, part of the explanation is likely due to the fact that the Seattle road network is much more densely connected than the US highway network, with more than twice the links per node (6.3 vs. 3.1). As a result, agents in Seattle have many more closely substitutable potential routes between home and work, reducing the impact of the improvement on each individual link of the network on the average cost of commuting.

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.

Data Availability Statement

The data underlying this article are available in the Zenodo repository and on the authors' webpages. The permanent link for the repository is https://doi.org/10.5281/zenodo.5608701.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online. And the replication packages are available at https://dx.doi.org/10.5281/zenodo.5608701.

A. APPENDIX: DERIVATIONS

The appendix presents the derivations of the results in Sections 3.1, 3.2, and 4.1; derivations for other sections are presented in Supplementary Appendix B.

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 4 by explicitly summing across all possible routes of all possible lengths. To do so, we sum across all locations that are travelled through all the possible paths as follows:

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

where k_n is the sub-index for the nth 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, that is:

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

where $\mathbf{A}^K=\left[A_{ij}^K\right]$, that is, 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 $\mathbf{B}=\left[b_{ij}\right]$ 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:

$$\pi_{ij}^{kl}=\sum_{r\in\Re_{ij}}\frac{\pi_{ij,r}}{\sum_{r'\in\Re_{ij}}\pi_{ij,r'}}n_r^{kl}\iff$$

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

$$\pi_{ij}^{kl} = \tau_{ij}^{\theta} \sum_{r \in \Re_{ij}} \prod_{l=1}^{K} t_{r_{l-1},r_l}^{-\theta} n_r^{kl},$$

where the second line used either equation (2) (for the economic geography model) or equation (12) (for the urban model), and the third line used the definition of τ_{ij} from equation (4).

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 $\mathbf{B} \equiv \left[b_{ij}\right] = \left[\tau_{ij}^{-\theta}\right]$ 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 $K-B-1$:

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

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:

$$\pi_{ij}^{kl} = \frac{b_{ik} a_{kl} b_{lj}}{b_{ij}} \iff$$

$$\pi_{ij}^{kl} = \frac{\tau_{ik}^{-\theta} t_{kl}^{-\theta} \tau_{lj}^{-\theta}}{\tau_{ij}^{-\theta}}, \tag{A.2}$$

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:

$$\Xi_{kl} = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \sum_{r \in \Re_{ij}} \pi_{ij,r} n_r^{kl} E_j \iff$$

$$\Xi_{kl} = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \pi_{ij}^{kl} X_{ij} \iff$$

$$\Xi_{kl} = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \frac{\tau_{ik}^{-\theta} t_{kl}^{-\theta} \tau_{lj}^{-\theta}}{\tau_{ij}^{-\theta}} \times \tau_{ij}^{-\theta} \frac{Y_i}{\Pi_i^{-\theta}} \frac{E_j}{P_j^{-\theta}} \iff$$

$$\Xi_{kl} = t_{kl}^{-\theta} \sum_{i \in \mathcal{N}} \tau_{ik}^{-\theta} \frac{Y_i}{\Pi_i^{-\theta}} \sum_{j \in \mathcal{N}} \tau_{lj}^{-\theta} \frac{E_j}{P_j^{-\theta}},$$

where the second line used equations (2) 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 (8) and (7)) 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:

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

$$\Xi_{kl}=\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}}\pi_{ij}^{kl}L_{ij}\iff$$

$$\Xi_{kl}=\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}}\frac{\tau_{ik}^{-\theta}t_{kl}^{-\theta}\tau_{lj}^{-\theta}}{\tau_{ij}^{-\theta}}\times\tau_{ij}^{-\theta}\frac{L_i^R}{\Pi_i^{-\theta}}\frac{L_j^F}{P_j^{-\theta}}\iff$$

$$\Xi_{kl}=\frac{\bar{L}}{W^\theta}\times t_{kl}^{-\theta}\times\left(\sum_{i\in\mathcal{N}}\tau_{ik}^{-\theta}\frac{L_i^R}{\Pi_i^{-\theta}}\right)\times\left(\sum_{j\in\mathcal{N}}\tau_{lj}^{-\theta}\frac{L_j^F}{P_j^{-\theta}}\right),$$

where the second line used equations (12) and (22), the third lines used equation (23), and the fourth lined rearranged. We substitute in the consumer and producer market access defined in (17) and (18) 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 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 article. Note that $\tau_{ij}^{-\theta}=[\mathbf{I}-\mathbf{A}]_{ij}^{-1}$, where $\mathbf{A}=\left[a_{ij}\right]=\left[t_{ij}^{-\theta}\right]$ is the adjacency matrix, so with a change of notation, we can rewrite the summation term as a matrix product:

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

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

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:

$$[\mathbf{I}-\mathbf{A}]\times\left[\bar{A}_i^{-\theta}y_i^{1+\theta}l_i^{-\theta(1+\alpha)}\right]=\frac{L^{(\alpha+\beta)\theta}}{W^\theta}\times\left[\bar{u}_i^\theta y_i^{1+\theta}l_i^{\theta(\beta-1)}\right]\iff$$

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

$$\bar{A}_i^{-\theta}y_i^{1+\theta}l_i^{-\theta(1+\alpha)}-\sum_j a_{ij}\bar{A}_j^{-\theta}y_j^{1+\theta}l_j^{-\theta(1+\alpha)}=\frac{L^{(\alpha+\beta)\theta}}{W^\theta}\bar{u}_i^\theta y_i^{1+\theta}l_i^{\theta(\beta-1)}.$$

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:

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

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

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

$$\left[\bar{u}_i^{-\theta}y_i^{-\theta}l_i^{\theta(1-\beta)}\right]\times\left[\mathbf{I}-\mathbf{A}^T\right]=\frac{\bar{L}^{(\alpha+\beta)\theta}}{W^\theta}\times\left[\bar{A}_i^\theta y_i^{-\theta}l_i^{\theta(\alpha+1)}\right]\iff$$

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

$$\bar{u}_i^{-\theta} y_i^{-\theta} l_i^{\theta(1-\beta)} - \sum_j a_{ji} \bar{u}_j^{-\theta} y_j^{-\theta} l_j^{\theta(1-\beta)} = \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}} \bar{A}_i^{\theta} y_i^{-\theta} l_i^{\theta(\alpha+1)}.$$

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

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

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

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 Supplementary Appendix C.1). 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:

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

$$\left(\frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}}\right)^{\frac{\theta\lambda}{1+\theta\lambda}}\sum_j\left(\bar{t}_{ji}\bar{L}^{\lambda}\right)^{-\frac{\theta}{1+\theta\lambda}}\bar{A}_i^{\theta\frac{\theta\lambda}{1+\theta\lambda}}\bar{u}_i^{\theta}\bar{u}_j^{-\frac{\theta}{1+\theta\lambda}}y_j^{-\frac{\theta}{1+\theta\lambda}}l_j^{\frac{\theta(1-\beta)}{1+\theta\lambda}}\iff$$

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

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:

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

$$\left[\bar{u}_i^{-\theta}\left(l_i^R\right)^{-\theta\beta+1}\right]=\frac{L^{(\alpha+\beta)\theta}}{W^{\theta}}\times[\mathbf{I}-\mathbf{A}]^{-1}\times\left[\bar{A}_j^{\theta}\left(l_j^F\right)^{\theta\alpha}\right]$$

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

$$\left[\bar{u}_i^{-\theta}\left(l_i^R\right)^{-\theta\beta+1}\right]-\mathbf{A}\times\left[\bar{u}_j^{-\theta}\left(l_j^R\right)^{-\theta\beta+1}\right]=\frac{L^{(\alpha+\beta)\theta}}{W^{\theta}}\times\left[\bar{A}_i^{\theta}\left(l_i^F\right)^{\theta\alpha}\right]\iff$$

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

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:

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

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

$$\left[\bar{A}_i^{-\theta}\left(l_i^F\right)^{-\theta\alpha+1}\right]=\frac{L^{(\alpha+\beta)\theta}}{W^{\theta}}\times\left[\bar{u}_j^{\theta}\left(l_j^R\right)^{\theta\beta}\right]\times\left[\mathbf{I}-\mathbf{A}^T\right]^{-1}\iff$$

$$\left[\bar{A}_i^{-\theta}\left(l_i^F\right)^{-\theta\alpha+1}\right]\times\left[\mathbf{I}-\mathbf{A}^T\right]=\frac{L^{(\alpha+\beta)\theta}}{W^{\theta}}\times\left[\bar{u}_i^{\theta}\left(l_i^R\right)^{\theta\beta}\right]\iff$$

$$\left[\bar{A}_i^{-\theta}\left(l_i^F\right)^{-\theta\alpha+1}\right]-\left[\bar{A}_j^{-\theta}\left(l_j^F\right)^{-\theta\alpha+1}\right]\times\mathbf{A}^T=\frac{L^{(\alpha+\beta)\theta}}{W^{\theta}}\times\left[\bar{u}_i^{\theta}\left(l_i^R\right)^{\theta\beta}\right]\iff$$

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

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 Supplementary Appendix C.1), incorporates endogenous traffic congestion. For the first equilibrium condition (30), we have:

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

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

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

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

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

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

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

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

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

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

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

as claimed.

REFERENCES

ADAO, R., ARKOLAKIS, C. and ESPOSITO, F. (2021), “General Equilibrium Effects in Space: Theory and Measurement” (mimeo).

AHLFELDT, G. M., REDDING, S. J., STURM, D. M. and WOLF, N. (2015), “The Economics of Density: Evidence from the Berlin Wall”, *Econometrica*, **83**, 2127–2189.

AKAMATSU, T. (1996), “Cyclic Flows, Markov Process and Stochastic Traffic Assignment”, *Transportation Research Part B: Methodological*, **30**, 369–386.

ALLEN, T. and ARKOLAKIS, C. (2014), “Trade and the Topography of the Spatial Economy”, *The Quarterly Journal of Economics*, **129**, 1085–1140.

ALLEN, T., ARKOLAKIS, C. and LI, X. (2020), "On the Equilibrium Properties of Network Models with Heterogeneous Agents" (NBER Working Paper, (w27837)).

ALLEN, T., ARKOLAKIS, C. and TAKAHASHI, Y. (2020), "Universal Gravity", *Journal of Political Economy*, **128**, 393–433.

ANDERSON, J. E., and VAN WINCOOP, E. (2003), "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review*, **93**, 170–192.

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

ASCE (2017), "2017 Infrastructure Report Card" https://www.infrastructurereportcard.org/wp-content/uploads/2016/10/2017-Infrastructure-Report-Card.pdf.

BECKMANN, M. J., MCGUIRE, C. B. and WINSTEN, C. B. (1955), "Studies in the Economics of Transportation" https://trid.trb.org/view/91120.

BELL, M. G. (1995), "Alternatives to Dial's Logit Assignment Algorithm", *Transportation Research Part B: Methodological*, **29**, 287–295.

CALIENDO, L., PARRO, F., ROSSI-HANSBERG, E. and SARTE, P.-D. (2018), "The Impact of Regional and Sectoral Productivity Changes on the US Economy", *The Review of Economic Studies*, **85**, 2042–2096.

CBO (1982), "The Interstate Highway System: Issues and Options" *Congressional Budget Office of the United States*. https://www.cbo.gov/sites/default/files/97th-congress-1981-1982/reports/doc19b-entire.pdf.

CHANEY, T. (2018), "The Gravity Equation in International Trade: An Explanation", *Journal of Political Economy*, **126**, 150–177.

CHARTRAND, G. (1977), *Introductory Graph Theory* (Courier Corporation).

COMBES, P.-P., and GOBILLON, L. (2015), "The Empirics of Agglomeration Economies", in Duranton, G., Henderson, J. V. and Strange, W. C. (eds) *Handbook of Regional and Urban Economics*, vol. 5 (Elsevier) 247–348.

COSTINOT, A., and RODRÍGUEZ-CLARE, A. (2014), "Trade Theory with Numbers: Quantifying the Consequences of Globalization", in Gopinath, G., Helpman, E. and Rogoff, K. (eds) *Handbook of International Economics*, vol. 4 (Elsevier) 197–261.

COUTURE, V., DURANTON, G. and TURNER, M. A. (2018), "Speed", *Review of Economics and Statistics*, **100**, 725–739.

DE PALMA, A., KILANI, M. and LINDSEY, R. (2005), "Congestion Pricing on a Road Network: A Study using the Dynamic Equilibrium Simulator METROPOLIS", *Transportation Research Part A: Policy and Practice*, **39**, 588–611.

DE PALMA, A., LINDSEY, R., QUINET, E. and VICKERMAN, R. (2011), *A Handbook of Transport Economics* (Edward Elgar Publishing).

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

DISDIER, A.-C., and HEAD, K. (2008), "The Puzzling Persistence of the Distance Effect on Bilateral Trade", *The Review of Economics and Statistics*, **90**, 37–48.

DONALDSON, D. (2015), "The Gains from Market Integration", *Economics*, **7**, 619–647.

——— (2018), "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure", *American Economic Review*, **108**, 899–934.

DONALDSON, D., and HORNBECK, R. (2016), "Railroads and American Economic Growth: A "Market Access" Approach", *The Quarterly Journal of Economics*, **131**, 799–858.

DUCRUET, C., JUHÁSZ, R., NAGY, D. K., STEINWENDER, C. ET AL. (2020), "All Aboard: The Effects of Port Development" (Discussion Paper).

DURANTON, G., and TURNER, M. A. (2011), "The Fundamental Law of Road Congestion: Evidence from US Cities", *American Economic Review*, **101**, 2616–2652.

EATON, J., and KORTUM, S. (2002), "Technology, Geography and Trade", *Econometrica*, **70**, 1741–1779.

EDWARDS, A. (2017), "U.S. Cities List" (Discussion Paper).

ELURU, N., PINJARI, A., GUO, J., SENER, I., SRINIVASAN, S., COPPERMAN, R. and BHAT, C. (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 SCHAAL, E. (2020), "Optimal Transport Networks in Spatial Equilibrium", *Econometrica*, **88**, 1411–1452.

FAN, J., LU, Y. and LUO, W. (2019), "Valuing Domestic Transport Infrastructure: A View from the Route Choice of Exporters", *The Review of Economics and Statistics*, 1–46.

FAN, J. and LUO, W. (2020), "A Tractable Model of Transshipment". https://fanjt.weebly.com/uploads/1/9/4/7/19473457/transshipment_note.pdf.

FEDERAL HIGHWAY ADMINISTRATION (2015), "2015 Status of the Nation's Highways, Bridges, and Transit: Conditions and Performance" (Discussion Paper, US Department of Transportation).

FEIGENBAUM, B., FIELDS, M. G. and PURNELL, S. (2020), "25th Annual Highway Report" (Reason Foundation).

FEYRER, J. (2019), "Trade and Income: Exploiting Time Series in Geography", *American Economic Journal: Applied Economics*, **11**, 1–35.

FHA (2008), "2008 Status of the Nation's Highways, Bridges, and Transit: Conditions and Performance" (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**, 163–197.

FOGEL, R. W. (1964), *Railroads and American Economic Growth: Essays in Econometric History*, vol. 296 (Baltimore: Johns Hopkins Press).
GALICHON, A. (2016), *Optimal Transport Methods in Economics* (Princeton University Press).
GANAPATI, S., WONG, W. F. and ZIV, O. (2020), "Entrepôt: Hubs, Scale, and Trade Costs" (Discussion Paper).
HEAD, K. and RIES, J. (2001), "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade", *American Economic Review*, **91**, 858–876.
HEBLICH, S., REDDING, S. J. and STURM, D. M. (2020), "The Making of the Modern Metropolis: Evidence from London", *The Quarterly Journal of Economics*, **135**, 2059–2133.
HILLBERRY, R. and HUMMELS, D. (2008), "Trade Responses to Geographic Frictions: A Decomposition using Micro-data", *European Economic Review*, **52**, 527–550.
HOLDEN, D. J. (1989), "Wardrop's Third Principle: Urban Traffic Congestion and Traffic Policy", *Journal of Transport Economics and Policy*, **23**, 239–262.
HUMMELS, D. L., and SCHAUR, G. (2013), "Time as a Trade Barrier", *American Economic Review*, **103**, 2935–2959.
KARLIN, S., and NIRENBERG, L. (1967), "On a Theorem of P. Nowosad", *Journal of Mathematical Analysis and Applications*, **17**, 61–67.
KASZYNSKI, W. (2000), *The American Highway: The History and Culture of Roads in the United States* (Jefferson, North Carolina, and London: McFarland & Company, Inc.).
KRUGMAN, P. (1991), "Increasing Returns and Economic Geography", *The Journal of Political Economy*, **99**, 483–499.
LEFEVRE, B., LEIPZIGER, D. and RAIFMAN, M. (2014), "The Trillion Dollar Question: Tracking Public and Private Investment in Transport" (Washington, DC: World Resources Institute).
LIND, N. and RAMONDO, N. (2018), "Trade with Correlation" (Discussion Paper, National Bureau of Economic Research).
MATTSSON, L.-G., WEIBULL, J. W. and LINDBERG, P. O. (2014), "Extreme Values, Invariance and Choice Probabilities", *Transportation Research Part B: Methodological*, **59**, 81–95.
MONTE, F., REDDING, S. J. and ROSSI-HANSBERG, E. (2018), "Commuting, Migration, and Local Employment Elasticities", *American Economic Review*, **108**, 3855–3890.
NSTIFC (2009), "Paying Our Way: A New Framework for Transportation Finance", *National Surface Transportation Infrastructure Financing Commission*. https://rosap.ntl.bts.gov/view/dot/17160.
OFFICE OF THE STATE AUDITOR, DIVISION OF TECHNICAL ASSISTANCE. (2002), *Capitalization and Depreciation of Infrastructure* (Mississippi Association of Governmental Purchasing and Property Agents).
OSHER, S. and SETHIAN, J. (1988), "Fronts Propagating with Curvature-dependent Speed: Algorithms based on Hamilton-Jacobi Formulations", *Journal of Computational Physics*, **79**, 12–49.
PASCALI, L. (2017), "The Wind of Change: Maritime Technology, Trade, and Economic Development", *American Economic Review*, **107**, 2821–54.
POLYANIN, A. and MANZHIROV, A. (2008), *Handbook of Integral Equations* (Chapman & Hall/CRC).
RAMONDO, N. and RODRÍGUEZ-CLARE, A. (2013), "Trade, Multinational Production, and the Gains from Openness", *Journal of Political Economy*, **121**, 273–322.
REDDING, S. and VENABLES, A. J. (2004), "Economic Geography and International Inequality", *Journal of International Economics*, **62**, 53–82.
REDDING, S. J. (2016), "Goods Trade, Factor Mobility and Welfare", *Journal of International Economics*, **101**, 148–167.
REDDING, S. J. and ROSSI-HANSBERG, E. (2017), "Quantitative Spatial Economics", *Annual Review of Economics*, **9**, 21–58.
REDDING, S. J. and TURNER, M. A. (2015), "Transportation Costs and the Spatial Organization of Economic Activity", in Duranton, G., Henderson, J. V. and Strange, W. C. (eds) *Handbook of Regional and Urban Economics*, vol. 5. (Elsevier) 1339–1398.
ROSENTHAL, S. S. and STRANGE, W. C. (2004), "Evidence on the Nature and Sources of Agglomeration Economies", in Henderson, J. V. and Thisse, J.-F. (eds) *Handbook of Regional and Urban Economics*, vol. 4. (Elsevier) 2119–2171.
SHEFFI, Y. (1985), *Urban Transportation Networks*, vol. 6 (Englewood Cliffs, NJ: Prentice-Hall).
SZABO, F. (2015), *The Linear Algebra Survival Guide: Illustrated with Mathematica* (San Diego, CA: Academic Press).
TSITSIKLIS, J. (1995), "Efficient Algorithms for Globally Optimal Trajectories", *IEEE Transactions on Automatic Control*, **40**, 1528–1538.
TSIVANIDIS, N. (2018), "Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio", https://static1.squarespace.com/static/55bb98e2e4b0ba843f39599e/t/5d039b14655edc0001d608a9/1560517406526/TsivanidisTransMilenio_6.2019.pdf.
VICKREY, W. (1967), "Optimization of Traffic and Facilities", *Journal of Transport Economics and Policy*, **1**, 123–136.
WEBER, H. J. and ARFKEN, G. B. (2003), *Essential Mathematical Methods for Physicists*. (San Diego, CA: Academic Press).
WEBER, S. and PÉCLAT, M. (2017), "GEOROUTE: Stata Module to Calculate Travel Distance and Travel Time between Two Addresses or Two Geographical Points" https://ideas.repec.org/c/boc/bocode/s458264.html.