Multimodal Transport Networks*

Simon Fuchs — Atlanta Fed

Woan Foong Wong — University of Oregon & CEPR

February 2025

Abstract

We examine the economic and environmental impacts of improvements and disruptions in multimodal transport networks. Our quantitative spatial equilibrium model incorporates routing over multiple modes and congestion at intermodal terminals. We estimate a modal substitution elasticity with highway and rail data, and a terminal congestion elasticity with vessel-positioning data. Calibrated to the U.S. freight network, our model identifies key bottlenecks and quantifies $300-700 million in additional real GDP gains from intermodal terminal improvements. These gains are 2.5 times higher without congestion, and substitution away from roads yield additional environmental benefits. Losing rail network access, factoring in modal substitution and general equilibrium effects, is estimated to reduce real GDP by $230 billion.

JEL Code: F11, R12, R42

Keywords: Multimodal transport, Transport network, Spatial equilibrium, Endogenous transport costs, Infrastructure investments, Disruptions, Bottlenecks

*Contact: simon.fuchs@atl.frb.org and wfwong@uoregon.edu. We are grateful to the NBER, DOT, and NSF for project support under the "Economics of Transportation in the 21st Century" Initiative. We thank Treb Allen, Adina Ardelean, Costas Arkolakis, Bruce Blonigen, Matilde Bombardini, Mark Colas, Arnaud Costinot, Doireann Fitzgerald, Cecile Gaubert, David Hummels, Samuel Kortum, Nelson Lind, Yuhei Miyauchi, Andrii Parkhomenko, Joseph Shapiro, Hugo E. Silva, Conor Walsh, and seminar participants at UC Berkeley Haas School of Business, Emory University, Kiel Institute, Minneapolis Federal Reserve Bank, National University of Singapore, Purdue University, UBC Sauder School of Business, and University of Stavanger, as well as participants at the NBER Spring 2023 Conference on Economics of Transportation, 1st Annual World Bank Conference on Transport Economics, ADB Urban Infrastructure conference, ASSA 2023, BFI Junior Spatial conference, CESifo Global Economy Conference, CRED Workshop on Urban and Regional Economics 2024, EIIT Conference (UC San Diego), Midwest International Trade Conference (GA Tech), Ron Jones Workshop (Rochester University), SMU-Jinan Conference on Urban and Regional Economics 2022, SED 2023, International Trade and Macro Workshop (St Louis Federal Reserve Bank), Trade Mini-Conference (UCLA), UEA, West Coast Trade Workshop (UC Davis), and West Coast Spatial Workshop (UC Berkeley) for very helpful comments. Benjamin Delgado and Philip Economides provided excellent research assistance. The views in this paper are solely the responsibility of the authors and should not necessarily be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System. All errors are our own.

Firms, workers, and locations are intrinsically linked via transport and trade networks. Improvements or disruptions within these networks can have large impacts on economic activity and welfare (Redding and Turner, 2015; Santamaria, 2020; Allen and Arkolakis, 2022). The current literature has mostly focused on just one mode of transport like road or ocean, but transportation takes place over multiple modes, driven largely by geography and containerization (Hummels, 2007; Coşar and Demir, 2018). The intermodal terminals within multimodal networks, enabling switching between modes, therefore are crucial for integrating the entire transport network, and congestion at these terminals can amplify disruptions or reduce the benefits from improvements. Furthermore, given that transportation is the largest source of direct US greenhouse gas emissions (EPA, 2024) and accounts for a quarter of global emissions (UN, 2021), these changes within the network have substantial environmental implications.

We study multimodal transport networks and their impact on the economic and environmental returns to infrastructure investments and disruptions. In particular, we focus on identifying and mitigating bottlenecks to improve the efficiency of the multimodal transport networks. Bottlenecks can arise from capacity constraints or limited alternatives, factors that are particularly relevant for intermodal terminals. Our analysis examines how bottlenecks are determined by the geography of the multimodal transport network, the location and congestion levels of intermodal terminals, as well as the relative cost of transport across modes. By incorporating these features, we can more accurately evaluate the overall impact of disruptions and infrastructure policies, including the importance of the rail network which was at risk of a nationwide shutdown due to the strike by rail workers in 2022.

We first develop a model that characterizes optimal routing across multiple transport modes in a tractable manner, and incorporate it into a quantitative spatial equilibrium framework. Goods traded from origin to destination via a specific route are a function of not just the origin being the least-cost producer (Eaton and Kortum, 2002) and the route being the least-cost option (Allen and Arkolakis, 2022), but also the route utilizing the least-cost transport mode available along each segment of the network. We account for the complexities of modeling multimodal transport networks by incorporating recent advances in transportation modeling and developing a model that recursively specifies optimal choices for sourcing, routes, *and* modes (Daly and Bierlaire, 2006).[1] Our framework has three main benefits. First, we can nest the choice of different transport modes within the optimal route choice model, with a distinct elasticity of substitution across modes. Second, we can include a rich set of mode-specific congestion patterns, including congestion at intermodal terminals. Third, we can derive closed-form expressions for equilibrium transport costs, traffic flows, and market access terms, which depend on mode and route choices as well as the topology of the transport network.

[1] Specifically, we embed a Nested Recursive Choice model that in principle allows for correlation between the stochastic shocks along the network and bypasses explicit solutions by deriving a recursive closed-form expression for transportation and sourcing costs, addressing the path enumeration problem without restricting agent choice or route distribution. In principle, the framework is flexible enough to capture topologically specific substitution patterns, allowing for a more accurate representation of network disruptions, localized congestion, and limited substitution patterns.

We formally demonstrate that changes within the multimodal transport network can generate both direct and indirect effects across different modes and throughout the entire network. For instance, improving the highway network will directly decrease road transport costs, which leads to an increase in the relative cost of other transport modes like rail, leading to a decline in their relative use to roads——a *modal substitution* effect. Simultaneously, better road access will enhance the market access of locations by reducing overall transport costs, thereby increasing their demand for transportation across all modes——a general equilibrium *complementarity* effect. The extent to which changes in transport costs increase market access can serve as a natural measure of how much of a bottleneck a particular segment is. More central segments with few alternatives naturally have the largest impact on market access across locations. Additionally, given that transport modes emit differing levels of emissions——trucks emit eight times more CO_2 per ton-mile than rail (CBO, 2022)——our framework can uniquely evaluate the environmental impacts stemming from both the substitution and complementarity effects.

In order to take our model to the data, we require two central parameters: the elasticity of substitution across modes and an estimate of congestion effects at intermodal terminals. We estimate the elasticity of modal substitution by studying how traffic shifts between modes in response to cost changes in a one mode, such as how relative rail and road traffic flows response to a decrease in road costs. By matching confidential waybill rail traffic data to truck traffic in cities, our approach combines the mode-specific cost change research design by Duranton and Turner (2011), which approximates reductions in road cost with increases in interstate highways, with the analysis on relative traffic shift between modes as in Hummels and Schaur (2013) and Lugovskyy, Skiba and Terner (2022). We find that a 1 percent increase in interstate highways result in a 0.9-1.2 percent decrease in relative rail to truck traffic use. Using the most conservative specification, we estimate the elasticity of modal substitution to be 1.099.

Next, we evaluate the impact of congestion at intermodal terminals using vessel positioning data down to the minute interval and measuring the time ships spend loading and unloading at ports. We estimate the elasticity of terminal congestion by investigating how responsive the ship's dwell times are to overall port traffic at the time of the ship's arrival. Since ship dwell times can be endogenous to conditions at port, we develop a shift-share instrument in order to identify the impact of overall port traffic on ship dwell times. We find that a 1 percent increase in port traffic increases ship dwell times by 0.24-0.26 percent. Utilizing the conversion established by Hummels and Schaur (2013), we estimate the elasticity of intermodal terminal congestion to be 0.096. While this elasticity is based on port data, we demonstrate its broader applicability to the entire multimodal network, including rail terminals, by highlighting a positive and significant relationship between railcar dwell times and port traffic.

To conduct counterfactuals, we first derive the counterfactual equilibrium, and then calibrate our framework to the US domestic transportation network. We derive the counterfactual equilibrium using 'hat algebra' (Dekle, Eaton and Kortum, 2008), calibrated with data on observed modal traffic, spatial distributions of income and labor densities, and key parameters such as the elasticity of modal

substitution and elasticity of congestion at intermodal terminals. The model structure allows us to describe this equilibrium through two interrelated systems of equations: the *aggregate equilibrium*, which calculates the spatial distribution of economic activity based on the endogenous transport costs, and the *transportation equilibrium*, which determines mode choice and congestion levels based on market access terms.[2] We then calibrate the model to fit and reproduce salient features of the US domestic transportation network, building a graph representation of the US multimodal transportation system drawing on high-resolution geo-spatial data on road, rail, and waterway linkages, as well as the location of intermodal terminals (both inland and ports). Using detailed traffic data for these modes, we apply our model to evaluate the economic and environmental impacts of targeted infrastructure investments and policy scenarios, taking into account the multimodal nature of the US domestic transport system.

Our main application assesses the welfare impact from improving the integration of the multimodal transport network, by investing in each intermodal terminal in the US. We find that terminals generating the largest gains are centrally located in the US, for example in Chicago and Dallas, highlighting the role of the multimodal network transporting goods from coastal to interior regions. Investments that lower transport costs by 1 percent in the most important terminals would generate an aggregate welfare gain equivalent to an additional $300-700 million of GDP (in 2012 USD). Congestion significantly affects these gains: welfare benefits are 2.5 times higher on average without congestion for the top terminals. These investments also result in modal substitution, shifting traffic use from road to rail and barges. Since trucks generate more greenhouse gases relative to trains, these substitution effects result in unintended environmental benefits valued at $23-45 million. These findings align with the goals of the inter-agency US National Blueprint for Transportation Decarbonatization, which emphasizes improving multimodal freight transport, and the European Green Deal, which aims to shift 75% of EU inland freight from road to rail and inland waterways (DOE, DOT, EPA, and HUD, 2024; European Commission, 2021).

We also quantify the welfare and environmental impacts of three policy scenarios. First, we evaluate the value of the US rail network by exploring the effects of the 2022 potential rail strike. While existing literature have quantified the significance of the road and highway transport network, less attention has been given to the value of the rail network apart from historical analyses. We find that losing access to the rail transport network would decrease US GDP by approximately $230 billion, inclusive of the mitigating effects of modal substitution towards roads and waterways, as well as the broader general equilibrium effects of reduced market access due to the loss of rail. This welfare loss is exacerbated by congestion effects, and the shift towards the more polluting road transport would cause additional environmental damage, with an estimated social cost of $12 billion.

Second, we consider the consequences of repealing the Jones Act from 1920, an active US trade regulation that restricts all ships transporting goods and passengers between US ports to be majority

[2] As noted by Galichon (2021), the solution to the (gravity trade) equilibrium system can be interpreted to be identical to the Sinkhorn algorithm - a common matrix scaling problem - and inherits its computational characterists both in terms of speed and scalability (Cuturi, 2013).

owned and crewed by US citizens, as well as built locally. By allowing foreign ships to operate on US domestic waterways, thereby lowering costs, we estimate that repealing the Jones Act would result in a welfare gain equivalent to increasing US GDP by about $3.2 billion. Our finding falls on the lower end of existing government estimates, likely due to our focus on the continental US and also the long-run adjustments of the multimodal network which has shifted away from domestic waterways due to higher costs and restrictive regulations. The repeal of the Jones Act would also allow for less reliance on trucks and rail, which results in a large environmental benefit since waterway transport is greener. Our third scenario investigates the impact of the Panama Canal disruptions due to drought. The welfare cost of decreased canal access is estimated to decreasing US GDP by about $2.7 billion. To the best of our knowledge, this is the first Panama Canal disruption estimate which accounts for both modal and route substitutions. We also find environmental losses due to the substitution towards both trucks and rail.

Related Literature. Our paper is related to a number of different strands of research. First, this paper contributes to a rapidly expanding literature incorporating realistic transportation network into quantitative spatial equilibrium models (see Redding (2020) for a recent survey). Within that literature, there have been multiple efforts to merge the disaggregated network structure of transportation infrastructure with a general equilibrium economic geography model (Fajgelbaum and Schaal, 2020; Allen and Arkolakis, 2022; Fajgelbaum et al., 2023). Particularly, Allen and Arkolakis (2022) proposed a tractable way of incorporating the optimal routing choice into an spatial equilibrium model in order to examine the general equilibrium implications of infrastructure improvements. While much theoretical progress has been made, the literature has often focused on one mode of transport—road, maritime, or rail (Coşar and Demir, 2018; Brancaccio, Kalouptsidi and Papageorgiou, 2020; Heiland et al., 2019; Ganapati, Wong and Ziv, 2021; Wong, 2022; Degiovanni and Yang, 2023; Dunn and Leibovici, 2023; Gibbons, Heblich and Pinchbeck, 2024). Focusing on roads, Fan, Lu and Luo (2019) and Jaworski, Kitchens and Nigai (2023) study how domestic road and highways affect transport costs and welfare in China and US respectively.[3] Our project adds to this literature by studying the general equilibrium analysis of the US multimodal transport system—highways, rail, and barges—as well as intermodal terminals which allow for switching across modes. Our model and elasticity estimates then allows us to quantify the economic and environmental impacts of infrastructure investments and policy scenarios, while accounting for the substitution and complementarity effects within the multimodal transport network.

A more recent literature has focused on the importance of ports and the gains from port development and efficiency. Brooks, Gendron-Carrier and Rua (2018) examines how the adoption of container technology at ports affect local activity while Ducruet et al. (2020) identifies how this technology adoption can crowd out economic activity by increasing land rents. Bonadio (2021) investigates how improvements in roads and ports have different distributional consequences locally and abroad. Focusing

[3]Additionally, Fan and Luo (2020) is a note which characterizes bilateral transport costs and their elasticities with respect to transshipment costs.

on bulk shipping, Brancaccio, Kalouptsidi and Papageorgiou (2024) evaluates the returns to infrastructure investment at ports in the presence of disruptions and congestion. Our paper adds to this literature by studying intermodal terminals, which includes ports, in the context of the multimodal transport network, showing that they are not just important for market access internationally but also domestically. Furthermore, we estimate an elasticity of congestion at intermodal terminals, and show that congestion can significantly influence and compound the returns from terminal improvements. Specifically, the benefits from improving top terminals are on average 2.5 times higher without congestion.

Our paper is also related to a long-standing literature in transportation that examines route and mode choice both empirically and theoretically (McFadden, Winston and Boersch-Supan, 1986; Rich, Kveiborg and Hansen, 2011; Beuthe, Jourquin and Urbain, 2014; Winston, 1981). The state-of-the-art in transportation studies solves high-dimensional traffic assignment problems algorithmically accounting for both disaggregated heterogeneity in modal and route choice (see Notteboom (2018) for a recent survey).[4] We employ similar tools to those recently developed in transportation studies, specifically the *stochastic user equilibrium* where routes and modes are chosen subject to a stochastic perception error. However, we go beyond the transportation literature by fully embedding the stochastic user equilibrium into a spatial general equilibrium framework, where input and output markets across space clear and factor and output prices are endogenously determined. This allows us to quantify the welfare effects of infrastructure improvements and disruptions within the multimodal transport network, taking into account both the modal substitution and general equilibrium complementarity effects.

Additionally, our paper is related to the recent literature on the environmental impacts of transportation. Most of this literature focuses on the link between international trade and greenhouse gas emissions via transportation, and changes to this relationship in response to environmental regulations (Shapiro, 2016; Mundaca, Strand and Young, 2021; Lugovskyy, Skiba and Terner, 2022) or trade policies (Cristea et al., 2013). We instead highlight how infrastructure investments and disruptions can have environmental consequences via the multimodal transport network. While some of these papers have found compositional shifts in transport mode use due to regulation and trade policy changes,[5] our quantitative general equilibrium framework allows us to distinguish between the substitution and complementarity effects. Additionally, our framework allows us to investigate how both of these effects contribute towards the environmental impacts of infrastructure improvements and disruptions.

The rest of the paper is structured as follows. Section 2 describes our data and the US multimodal

[4]For a recent theoretical contribution compare Kitthamkesorn, Chen and Xu (2015) which solves for the traffic assignment problem allowing for both endogenous route and mode choice. A recent applied quantitative contribution in this literature models the multimodal linkages between the US and China with endogenous route choice and congestion at port locations (Li, Xie and Bao, 2022).

[5]Cristea et al. (2013) finds that trade liberalization between countries will increase trade from more distant partner countries, resulting in a proportional increase in air transport use and greenhouse gas emission. Lugovskyy, Skiba and Terner (2022) finds that environmental regulations capping CO_2 emissions from maritime shipping will substitute demand towards air transport, increasing total transport-related CO_2 emissions.

transport network. We detail the multimodal routing model in Section 3 and explain how we estimate our modal substitution and congestion elasticities in Section 4. Section 5 discusses how we incorporate our elasticities and calibrate our model to the US multimodal transport network, as well as provides a validity check of our framework. We apply our model to evaluate the welfare and environmental impacts of infrastructure investments and policy scenarios in Section 6 and conclude in Section 7.

2 Data: US Domestic Freight Transportation

In this section, we introduce our data sources and provide an overview of the US domestic transport system. As mentioned above, the movement of goods takes place over multiple modes of transportation. Figure 1 plots the value share of freight moved by different modes across various distances within the US. Trucks primarily handle freight for shorter distances below 100 miles, accounting for 90 percent of value (Figure 1), highlighting the role of the denser road network in facilitating transportation at the start and end of the movement of goods. This is commonly known as the first and last mile in freight transportation (Rodrigue, 2020; Ranieri et al., 2018). Over longer distances, however, freight is increasingly transported using multiple modes. For freight transported over 1,500 miles, multimodal transport accounts for more than half by value and more than 70% by weight (Figures 1 and A.1 respectively). For context, the road distance between Los Angeles and Chicago is about 2000 miles.

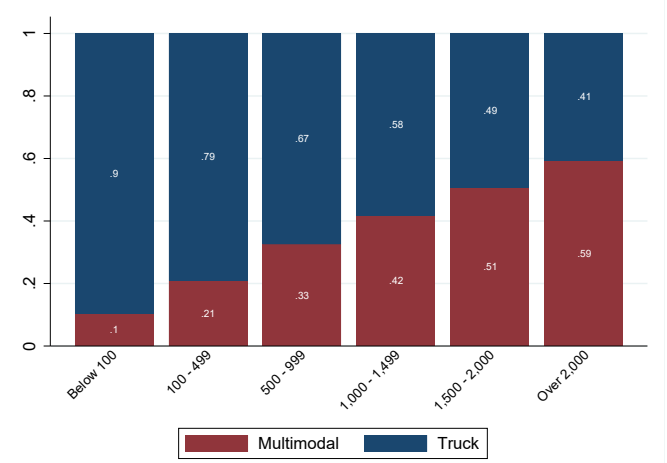
2.1 Rail, Road, Waterways, Ports, and Intermodal Terminals

Rail Geography Network and Traffic. We use detailed geo-spatial information from the Topologically Integrated Geographic Encoding and Referencing Database (TIGER, Census Bureau) to construct a graph representation of the US intermodal rail network. To do so, we subset the original network to segments that are owned by the largest Class I carriers and that are compatible with multimodal transport.[6]

For rail traffic data, we have obtained access to confidential carload waybills data from the Surface Transportation Board (see Section B.1.1 for further details). The waybill data captures the origin, destination, and interchange rail stations for freight cargo transported through the rail network. This detailed geographical information for rail cargo journeys allows us to study their routes. Additionally, the data set contains information on number of car loads, weight, and intermodality. Intermodality indicates if rail freight movement involved other transport modes, which is almost entirely containers. The proportion of freight transported over multiple modes is increasing dramatically over time: intermodal container rail traffic is the fastest growing rail traffic segment, having grown by more than 5 times since

[6]Class I railroads are the largest carriers operating on the US railroad system. They were originally in 1992 defined to be those carriers above $250m dollars of revenue, a cutoff that has been adjusted for inflation since. In 2021 the threshold stands at approximately $943m. There are currently seven class I carriers and they make up the large majority of the domestic rail freight market.

Figure 1. US Transport Mode Value Shares by Distance



Notes: This figure plots the observed value share of cargo transported by different modes across various distances in 2018. Multimodal indicates cargo movement that involves more than one mode, including truck, rail, and waterways. Source: Freight Analysis Framework, US Department of Transportation, and authors' calculations.

1984 (Figure A.2). We restrict our rail traffic sample to intermodal freight transported by Class I carriers.

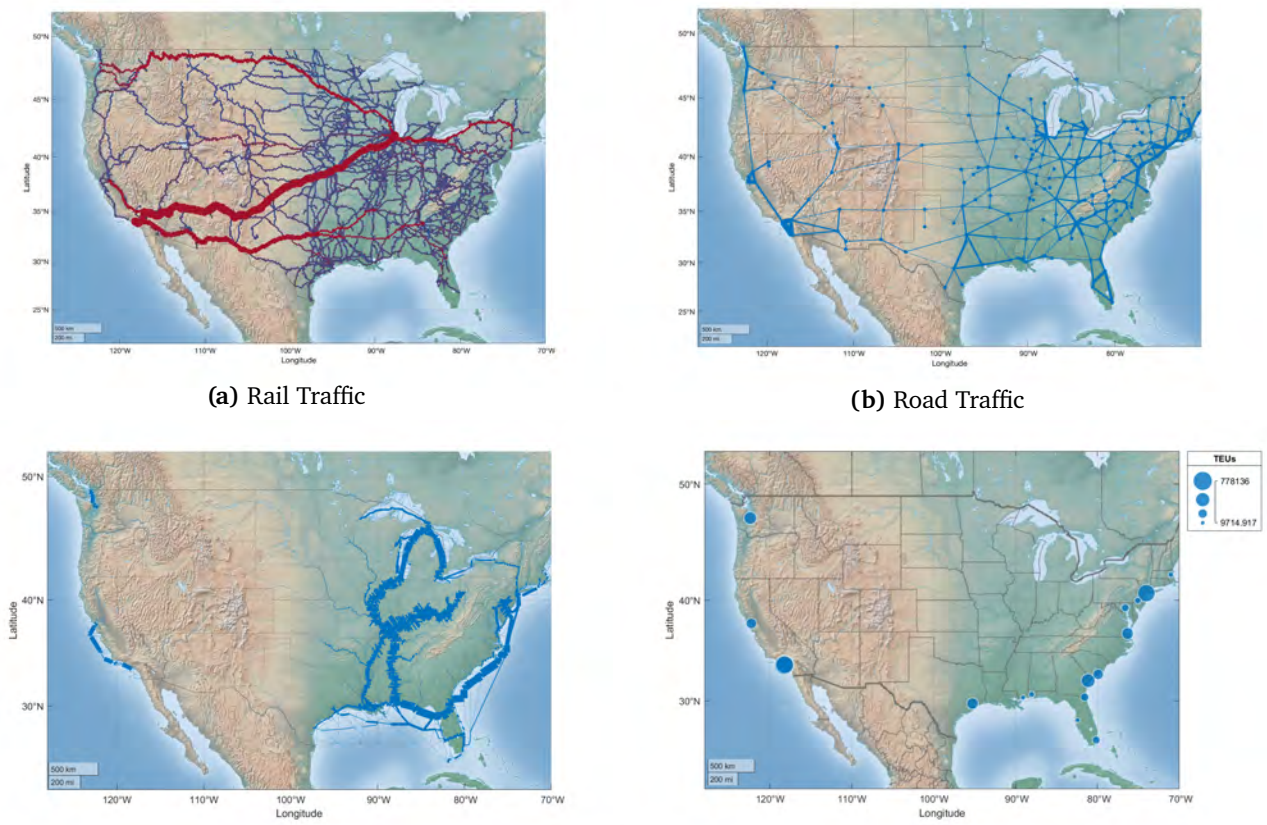
Figure 2a present the US rail intermodal traffic flows. Thicker lines, indicating higher traffic flows, link the US West and East coastal regions to interior locations like Chicago, Dallas, and Atlanta. Examples of such routes include Los Angeles and Chicago, New York City and Chicago, as well as Los Angeles and Dallas. Echoing Figure 1, the rail network serves to transport freight over long distances.

Road Geography Network and Traffic. We follow Allen and Arkolakis (2022), hereby abbreviated to AA2022, for the construction of road geo-spatial and traffic flows data (see Section B.1.2 for details). For road traffic data, the average annual daily traffic from the 2012 Highway Performance Monitoring System (HPMS) dataset by the Federal Highway Administration is used and the flows are allocated to individual links by constructing a length-weighted average of the annual daily traffic. Figure 2b shows a graph representation of the interstate highway system with thicker lines indicating higher road traffic flows. The highway flow patterns are quite different from rail (Figure 2a). Here, heavier highway traffic connects large densely populated cities that are either on the coast or in the interior regions, like Los Angeles and San Diego, Boston to Philadelphia, and the surrounding areas around Chicago.

Waterway Traffic. We next capture goods that are transported via US waterways by bringing in waterborne traffic data and location of inland ports from the US Army Corps of Engineers (USACE). We restrict our sample to manufactured goods which are primarily transported by barges. Figure 2c shows that most barge traffic is concentrated on the Great Lakes, Illinois River, and Mississippi River..

Port Location and Traffic. We include the top US container ports which are important coastal intermodal terminals facilitating the transfer and distribution of goods between ships and trucks or rail (see Section B.1.3 for details). The geospatial data on port location is obtained from the US Army Corps of Engineers. Container volumes at these ports come from the Port Performance Freight Statistics Program maintained by the Department of Transportation (DOT). Figure 2d visualizes the coastal ports in a

Figure 2. Traffic along the US Domestic Freight Transport System



(a) Rail Traffic

(b) Road Traffic

(c) Waterway Traffic

(d) Port Traffic

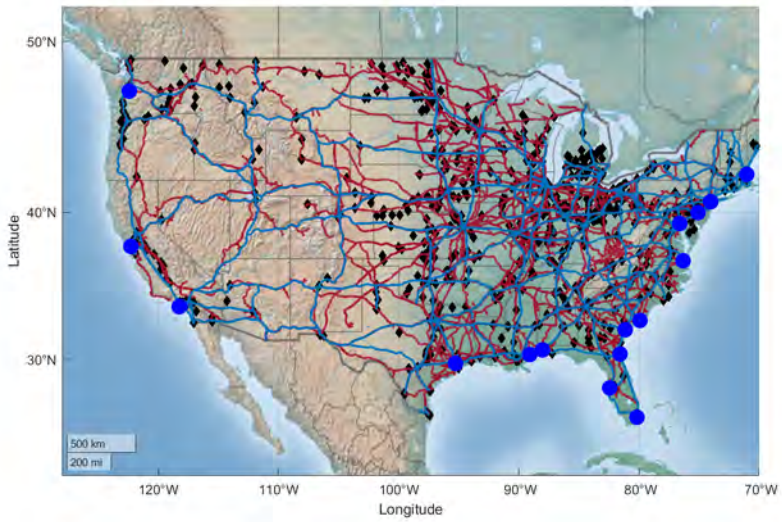
Notes: Panel (a) presents the US domestic rail traffic data for Class I carriers (the largest rail carriers) conditional on intermodal capability. Shortest routes are imputed between origin, interchange stations, and destination to assign total tonnage to individual rail segments along network. Thicker lines indicate higher traffic flows. Panel (b) presents the traffic along the graph representation of the interstate highway system. Thicker lines indicate higher traffic flows. Panel (c) presents the domestic waterborne traffic data for manufactured goods, where the shortest routes are imputed between origin and destination to assign total tonnage to individual segments of the domestic water network. Thicker lines indicate higher traffic flows. Panel (d) presents the container traffic volume at international ports. Larger circles indicate higher container volumes. Sources: Authors calculations, Confidential Carload Waybill, Surface Transportation Board; Highway Performance Monitoring System, the Federal Highway Administration; Waterborne Commerce statistics, US Army Corps of Engineers (USACE); and Port Performance Freight Statistics Program, Department of Transportation (DOT).

geographic bubble map, where the largest container ports include Los Angeles/Long Beach and New York (Newark). Between the coastal and inland ports, we capture 98 percent of total US container volumes.

Intermodal Terminals. A number of intermodal terminals are inland transfer points for goods moving between rail, road, and waterways. We capture the location of all intermodal terminals using the National Transportation Atlas Database (NTAD) maintained by the Department of Transportation (DOT).

US Multimodal Transport Network. To construct the US multimodal transport network, we start by combining all the geo-spatial information on each of the rail, road, and waterway networks. On top of that, we add the locations of intermodal terminals which includes coastal and inland ports, as well as inland intermodal terminals. The resulting graph representation of the US multimodal transport network consists of 288 nodes and 704 edges, and is presented in Figure 3. The nodes are either core-based statistical areas (CBSAs), or endpoints and intersections in the network.

Figure 3. US Multimodal Transportation Network



Notes: This figure shows the combined US multimodal freight network. The original GIS information is obtained from the U.S. Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) Database. The red lines indicate the Class I multimodal railroad network. The blue lines indicate the interstate highway system (IHS). Black diamonds indicate freight terminals that are owned by Class I operators allow for road-to-rail or rail-to-road intermodal movements. The blue circles indicate the top 18 ports.

2.2 Automatic Identification System (AIS) Vessel Traffic Data

We utilize automatic identification system (AIS) vessel traffic data from Marine Cadastre, a joint initiative between the Bureau of Ocean Energy Management and the National Oceanic and Atmospheric Administration. This data captures vessel location in US waters at 1-minute intervals using 200 land-based receiving stations. We observe the vessel's unique identifying International Maritime Organization Vessel number (IMO), its longitude and latitude location down to the minute, speed, and navigation status. The vessel's navigation status captures whether the vessel is being propelled (under way using engine), or moored—held in position at a pier. Using information on the ship's speed and navigation status, we define a ship's dwell time to be the time it spends being moored at a pier and has zero speed. This is a conservative measure of ship dwell time at ports because (1) a ship will spend time navigating within the port area as it prepares to moor at a pier and (2) a ship can also end up waiting outside of the port area at anchor before navigating to the port (New York Times, 2021).

In order to match these ships to the ports they are located at, we require geographical information of the ports. We use the Port Statistical Area shapefiles from the US Army Corps of Engineers and match these ships to the top 30 US container ports. These port polygon areas also allow us to calculate the total amount of time a ship spends within the port region on top of the time it spends moored at a dock. Additionally, in order to identify the cargo capacity of these ships and their containership status, we match these ships to the Port Entrance and Clearance dataset from the US Army Corps of Engineers using their identifying information and when they are at these ports. See Section B.1.4 for examples illustrating how we observe these ships and the duration of their stays at ports. The ship cargo capacity measures the volume of the ship that can be used for loading cargo (net tonnage of a ship). This cargo

capacity measure for each ship will contribute to our port traffic measure at each port every day.

Port Traffic Our measure of port traffic is defined as the sum of the net tonnage of each ship moored at the port each day, multiplied by the percent of the day they spend at the port—crucially including ships that arrived prior to that day but still remained moored at port. To be more specific, if a ship remained moored at port all the way without exiting, its contribution to port traffic would be 100% of their net tonnage (100% of the time it spent at the port). If a ship left at any point during that day, its net tonnage contribution would be less than 100% and be instead determined by the amount of time it spent moored at port that day. With this daily port traffic measure, we calculate moving averages of the port-level traffic for varying amounts of time. We have done this for 3, 7, 14, 21, and 28 days. We present the 28-day moving average results and have included the rest in the Online Appendix B.4.

Ship Dwell Times Summary Statistics Our matched dataset has 3,755 unique vessels with 1,444 containerships. The top 30 ports in our dataset account for around 95% of all US container trade by value and 98% by volume annually. For illustrative purposes, Figure 4a aggregates our minute- and ship-level dataset and plots the average of monthly containership dwell times across the top 30 US ports from June 2015 to December 2021. The average dwell time over this period is around 33.3 hours per ship with a standard deviation of 5 hours.[7] However, as seen in Figure 4a, there is a significant increase in the ship dwell times post 2021. The average ship dwell time after 2021 is 42.8 hours.

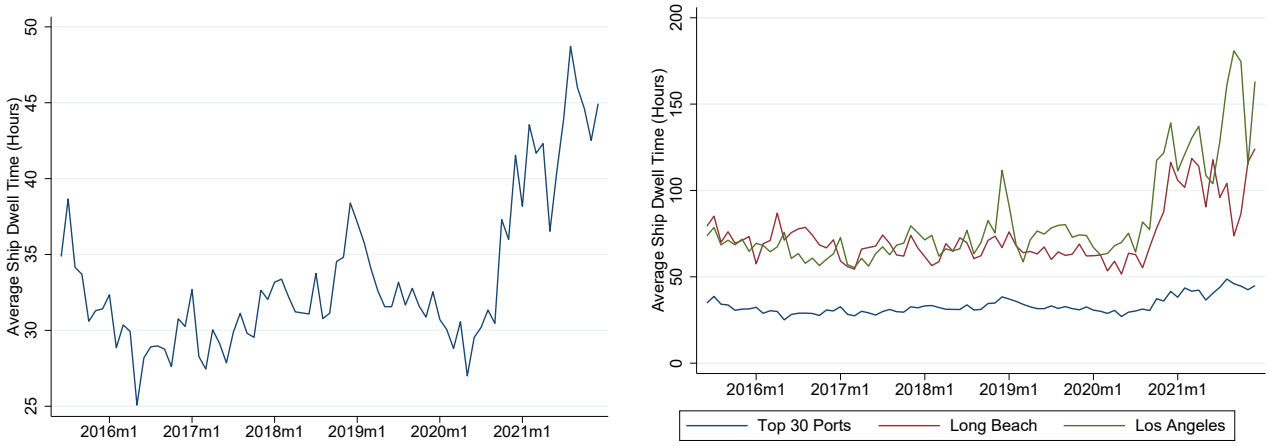
Additionally, there is considerable variation in ship dwell times across different ports and time periods. Figure 4b compares the average dwell times at the ports of Long Beach and Los Angeles (LA) to the overall average from Figure 4a. The dwell times at both ports are more than twice the overall average, 73.6 hours and 82.1 hours respectively. After 2021, dwell times at these ports surged even more dramatically, increasing to over 4 days at Long Beach (104 hours) and almost 6 days at LA (136 hours). These substantial variations highlight the growing congestion and delays at key intermodal terminals and the underlying interaction between ships and ports that contributes to this congestion.

3 Economic Geography Model with Multimodal Routing

In this section, we present a tractable model that characterizes optimal routing across multiple transport modes, while allowing for mode-specific congestion particularly at intermodal terminals. This model is then integrated into a standard economic geography model with domestic trade between discrete locations and freely mobile labor reallocation across locations (Allen and Arkolakis, 2014; Redding, 2016). When embedding optimal route choice across multiple modes into a spatial equilibrium model using a graph representation of the multimodal network, we encounter two technical challenges. First is

[7] Our averages are lower than the estimates in Brancaccio, Kalouptsidi and Papageorgiou (2024), in part due to their inclusion of wait time at anchorage and their focus on bulk ships.

Figure 4. Containership Dwell Times at Port



(a) Overall Average at top 30 Ports **(b)** Long Beach and Los Angeles vs Overall Average

Notes: Panel (a) plots the overall average of monthly containership dwell times across the top 30 US ports from June 2015 to December 2021. Panel (b) plots the monthly containership dwell times at the top two largest US ports, Long Beach and Los Angeles (red and green lines respectively), and compares their dwell times to the same Panel (a) overall average (blue line). Both panels are weighted by ship net tonnage. Sources: Authors calculations, Marine Cadastre, Bureau of Ocean Energy Management and the National Oceanic and Atmospheric Administration; International Maritime Organization; and Port Statistical Area, US Army Corps of Engineers (USACE).

the that the number of feasible routes can be overwhelming already in low-dimensional graphs, and this complexity is further increased in our case with the inclusion of multiple transport modes. This curse of dimensionality complicates path enumeration and makes brute-force solutions impractical. A second challenge is that existing approaches, which rely on invertible matrix expressions for cost-minimizing route choices (as in AA2022), do not readily accommodate mode choice where the mode choice elasticity differs from the route choice elasticity. We address both these challenges by formulating the multimodal route choice problem recursively with a nested mode choice at each segment along the route.[8]

Our modeling approach is outlined in six steps. First, Subsection 3.1 defines the multi-layered graph that represents the multimodal transport network, and specifies consumption and production. Second, Subsection 3.2 introduces the recursive routing choice. Subsection 3.3 then expands on this to include the combined routing and sourcing choice, and resulting equilibrium equations. Next, Subsection 3.4 presents the nested transport mode choice while Subsection 3.5 introduces congestion and the gravity equations for traffic flows. Finally, Subsection 3.6 demonstrates the substitution and complementarity effects within our model by examining the impact of a mode-specific cost change.

3.1 Setup

Geography and transport. Let $\mathcal{G} \equiv (\mathcal{N}, \mathcal{L})$ be a multigraph representing a multimodal transport network where $\mathcal{N}$ and $\mathcal{L}$ are the set of nodes and links respectively. We define the set of successor nodes

[8] One alternative method to introduce mode choice is to simply allow for mode-specific subnetworks. This naturally gives rise to a setting where mode choice is then *implied* by route choice, which represents a direct extension of AA2022 to transport mode choice (see Appendix A.5 for details). A previous version of this paper presents such a model. The limitation of this modeling approach is that it assumes that the elasticity of modal substitution must be identical to the elasticity of route choice. However, as we demonstrate in the next section, our estimates indicate that this assumption does not hold.

$\mathscr{F}(i)$ and the set of predecessor nodes $\mathscr{B}(i)$ for each node $i \in \mathscr{N}$. Furthermore, let $\mathscr{G}_m \equiv (\mathscr{N}_m, \mathscr{L}_m)$ be the subgraph representing the transport network for mode m, where $\mathscr{N}_m \subseteq \mathscr{N}$ and $\mathscr{L}_m \subseteq \mathscr{L}$ are the mode-specific nodes and links respectively. We also define the mode-specific successor and predecessor nodes $\mathscr{F}_m(i)$ and $\mathscr{B}_m(i)$. Each link ij in the set of m mode-specific links, $ij \in \mathscr{L}_m$, is associated with a generalized link travel cost $t_{ij,m}$ which is specific to mode m and can be endogenously dependent on flow. Within this multimodal transport network, we define a primary network, $m = 1$, that represents the dense road network on which all cities and road intersections are located. All other modes, including rail and barges, are secondary transport networks ($m \neq 1$). A subset of nodes on the primary and secondary networks are intermodal terminals which allow for switches between mode-specific networks.

Consumption and Production. A representative agent lives in location j, supplies her unit endowment of labor inelastically, earns a wage rate w_j, and purchases quantities of a continuum of consumption goods, $v \in [0,1]$. She is endowed with constant elasticity of substitution (CES) preferences where the elasticity of substitution is given by $\sigma \geq 0$. We define aggregate income as Y^W, total labor endowment as $\bar{L}$, and average per capita income as the numeraire, i.e. $Y^W/\bar{L} = 1$.

Each location i produces a good $v \in [0,1]$ subject to a constant returns to scale technology and transports it to destination j along a feasible route. As we detail below, transport costs are route-specific in the sense that they depend on which neighboring node the good is being sourced from. Assuming perfect competition, the price of good v in destination j from origin i along neighboring node k is given by $p_{ikj} = \frac{t_{kj}\tau_{ik}}{A_i} w_i$, where the marginal cost of production in i is $\frac{w_i}{A_i}$, local wages are w_i, each worker can produce A_i units of goods, and trade cost is route-specific.[9]

In what follows next, we will introduce recursive routing and - eventually - sourcing choice that further specifies how trade cost is being endogenously determined as a function of the underlying network and the agent's routing and modal choices.

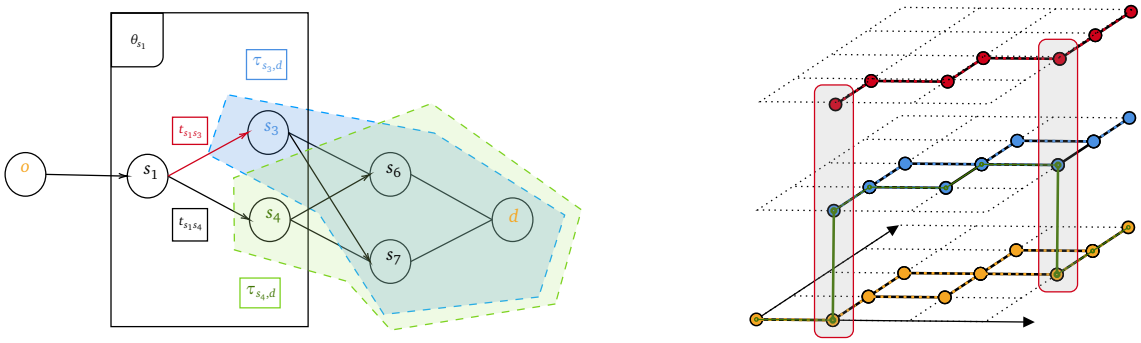
3.2 Recursive Routing Choice

We begin by presenting the recursive routing model before we extend the framework in the next subsection to encompass both routing and sourcing, building on recent advancements in discrete route choice modeling (Daly and Bierlaire, 2006; Melo, 2012; Oyama, Hara and Akamatsu, 2022).

The individuals' routing decision is characterized as a joint choice of elemental links within the route

[9] In this paper, we assume perfect competition among freight transport companies. While this assumption simplifies our analysis, it also aligns with our focus on multimodal transport networks. Literature shows that multimodal container transport is generally more competitive than unimodal transport and that rail transport of containers is more competitive than non-containerized shipments. Zgonc, Tekavčič and Jakšič (2019) finds that multimodal road-rail transport remains competitive over short distances compared to unimodal road transport. Similarly, Surface Transportation Board (2009) reports lower markups for multimodal rail shipments and lower rail rates near alternative transport modes. Detailed analysis of market power in container freight transport would require comprehensive door-to-door shipment data, which is currently unavailable.

Figure 5. Recursive and Multimodal Routing



(a) Recursive Routing on a Directed Graph

(b) Multimodal Routing on a Multi-layered Graph

Notes: Panel (a) shows a directed transport network from origin o to destination d, with nodes as circles and edges representing transport costs. At node s_1, an agent chooses between neighboring nodes s_4 and s_3, each having a continuation value for minimizing transport cost to d. Panel (b) depicts multimodal routing on a multi-layered graph, similar to Panel (a), with nodes as circles and edges linked by mode-specific transport costs. Three mode-specific graphs are shown: primary roads (yellow), secondary rail and barges (blue and red). The overall figure shows a multi-layered graph with a green example route and intermodal transfers highlighted at nodes where mode choice occurs.

in an ordered sequence from origin i toward destination j (Papola and Marzano, 2013). In modelling this joint choice recursively, the individual makes a cost-minimizing routing choice at each node by comparing the cost of accessing one neighboring node, which includes its link cost and the cost of continuing from there to the final destination, against the cost to access other neighboring nodes which includes their link costs and own subsequent continuing costs to the destination. Figure 5a illustrates an example of this. At node s_1, the individual makes a cost-minimizing choice between nodes s_3 and s_4 by comparing (1) the total cost to access node s_3, which includes its link cost ($t_{s_1s_3}$) and the continuing cost to destination d (τ_{s_3d}), against (2) the total cost to access node s_4, which includes its own link cost ($t_{s_1s_4}$) and the continuing cost to d (τ_{s_4d}).

Generalizing this, the individual makes a cost-minimizing routing choice at each node i by comparing the link cost of moving towards any neighboring node k (t_{ik}) and the cost of continuing from there to destination j (τ_{kj}). We further assume that the trade cost is route-specific and multiplicative over all links, and that the agent experiences a segment-specific preference shock, ε_{kj} is independently identically Frechet distributed across (successor) neighboring nodes $k \in \mathscr{F}(i)$ with a dispersion parameter θ_i that can be node-specific.[10] Exploiting the properties of the Frechet distribution, we derive a closed-form

[10]While we present the model using a sequentially drawn Frechet cost shock, it can alternatively be described as a generalized extreme value (GEV) type model where with recursively related Frechet shocks along the network topology. The GEV family of models relaxes independent error draws and allows for more generalized substitution relationships among alternatives (McFadden, 1977) by introducing a correlation function $G(\cdot)$ and $P_i = \frac{Y_i G_i}{G}$ being the choice probability of alternative i, where Y_i is the choice-specific return and G_i the partial derivative of G with regard to i. This model's isomorphic description, sometimes termed a Network-GEV model as introduced in Daly and Bierlaire (2006), uses the network's observed topological structure to provide a nesting structure for the choice model. Specifically, the Network-GEV model is defined by a correlation function $G(\cdot)$ that is recursively defined as, $G_{ij}^d = t_{ij}^{-\theta_i} \sum_{k \in \mathscr{F}(j)} \left(G_{jk}^d\right)^{\theta_i/\theta_k}$ if $j \neq d$, and $G_{ij}^d = t_{ij}^{-\theta_i}$ if $j = d$. Furthermore, using results from Lind and Ramondo (2023) this routing choice can be extended to define a sourcing choice where the correlational structure is defined by the observed topology of the transport network.

expression for the expected transport cost from i to j as below,

$$\tau_{ij} = \mathrm{E}\left[\min_{k\in\mathscr{F}(i)}\left\{\frac{t_{ik}\tau_{kj}}{\varepsilon_{kj}}\right\}\right] \propto \left(\sum_{k\in\mathscr{F}(i)}\left(t_{ik}\tau_{kj}\right)^{-\theta_i}\right)^{-\frac{1}{\theta_i}} \tag{1}$$

which provides a recursive relationship between transport costs along the network topology.[11] We highlight three characteristics of this framework. First, the definition of the transport cost explicitly relies on the topology of the network, as the agent's choice set at node i explicitly depends on the set of (successor) neighboring nodes, $\mathscr{F}(i)$, and the link costs associated with traversing the connecting edges. In our Figure 5a example, there are two neighboring nodes (s_3 and s_4). Second, this formulation allows for varying dispersion elasticities along the network topology (θ_i), allowing for a more flexible route and sourcing characterization than the isoelastic case. Unlike existing models that rely on IID technology or preference shocks and assume the independence of irrelevant alternatives (IIA), this approach can have more flexible substitution patterns. Third, with homogenous dispersion elasticities ($\theta_i = \theta$) and without considering multiple modes, this characterization is isomorphic to the routing characterization in AA2022,[12] while still retaining its analytically convenient feature.

3.3 Combined Routing and Sourcing Choice, and Recursive Equilibrium

We now combine the recursive routing choice in Equation (1) with the stochastic sourcing choice from Eaton and Kortum (2002) to derive a closed-form characterization for both route and sourcing choices. Next, we integrate the route and sourcing choice into the equilibrium conditions, relating endogenous equilibrium outcomes recursively along the network topology.

Routing and Sourcing. Under perfect competition, consumers (recursively) source their products by choosing the lowest-cost route-source combination. The price the consumer in j faces is the expected minimum of choosing a variety ν that is both the least cost source i and the being routed along the lowest cost route from i to j utilizing the neighboring node k,

$$p_j \equiv \mathbf{E}_\varepsilon\left[\min_{(k,i)\in\mathscr{B}(j)\times N}\frac{t_{kj}\tau_{ik}}{\varepsilon_{ikj}(\nu)}w_i\right] \tag{2}$$

where, following Eaton and Kortum (2002), we assume that $\varepsilon_{ikj}(\nu)$ is iid Fréchet distributed across the set of (predecessor) neighboring nodes $k \in \mathscr{B}(j)$ and sourcing partners i with scale parameter $1/A_i$,

[11] Here, we characterize the routing in terms of a forward moving routing choice along the successor nodes $k \in \mathscr{F}(i)$. Instead, one could also describe the same routing choice equivalently as a sourcing choice along the predecessor nodes $i \in \mathscr{B}(k)$. For $\theta_i = \theta$, it turns out that the two are equivalent, as discussed in Appendix A.5.

[12] This isomorphism between one-shot decisions over the universe of paths and sequential decision-making relying on dynamic programming has been previously established, most notably by Antràs and Gortari (2020) who explore a one-shot sequential sourcing decision in their main text and introduce a dynamic programming formulation in their Appendix A.1.3. Our work extends this formulation in two distinct ways: by explicitly formulating the decision-making on a graph, and by embedding the solution into the equilibrium conditions for a computationally convenient representation of the spatial equilibrium. See Appendix A.5 for further details.

where A_i captures origin-specific efficiency and θ_j is a node-specific shape parameter which regulates the inverse of shock dispersion.[13] Given the distribution of shocks, the probability that j sources a good from i that is routed via neighboring node k is as follows (see detailed derivations in Appendix A.1),

$$\underbrace{\pi_{ij,k}}_{\substack{\text{Joint routing-sourcing prob.}\\ \text{from } i \text{ to } j \text{ via node } k}} = \underbrace{\frac{\tau_{ij}^{-\theta_j}\left(\frac{w_i}{A_i}\right)^{-\theta_j}}{\sum_{i'\in N}\tau_{i'j}^{-\theta_j}\left(\frac{w_{i'}}{A_{i'}}\right)^{-\theta_j}}}_{\text{Sourcing prob. from } i} \times \underbrace{\frac{\left(t_{ik}\tau_{kj}\right)^{-\theta_j}}{\sum_{k'\in\mathscr{B}(j)}\left(t_{ik'}\tau_{k'j}\right)^{-\theta_j}}}_{\text{Routing prob. via node } k} \equiv \pi_{ij}\times\pi_{ij}^{ik} \tag{3}$$

where we have decomposed the combined routing-sourcing probability into two different components. The first term, π_{ij}, represents the sourcing probability given the expected bilateral transport cost, τ_{ij}, defined as and is identical to the sourcing probability in Eaton and Kortum (2002) except for the inclusion of heterogeneous substitution elasticities along the network θ_j. The second term, π_{ij}^{ik}, represents an implicit route choice probability, characterizing the likelihood of choosing a neighboring node k amongst all neighboring nodes when sourcing from origin i.

Recursive Equilibrium. To obtain the equilibrium conditions, we use our recursive sourcing-route probability from Equation (3) to derive a generalized gravity equation (see detailed derivations in Appendix A.2), then we impose two market clearing conditions. First, good markets clear—total income in location i, Y_i, is equal to its total sales ($Y_i = \sum_{j=1}^{N} X_{ij}$). Second, trade is balanced—total consumption expenditure in location j, E_j, is equal to its total imports ($E_j = \sum_{i=1}^{N} X_{ij}$). We obtain the following equilibrium conditions (see Appendix Section A.3 for details):

$$\Pi_i^{-\theta_i} = \left(t_{ii}^{-\theta_i}\right)\frac{\delta_i}{P_i^{-\theta_i}} + \sum_{k\in\mathscr{F}(i)} t_{ik}^{-\theta_i}\tilde{\Pi}_k^{-\theta_i}, \qquad P_j^{-\theta_j} = \left(t_{jj}^{-\theta_j}\right)\frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{k\in\mathscr{B}(j)} t_{kj}^{-\theta_j}P_k^{-\theta_j}, \tag{4}$$

where the equilibrium conditions solve for the endogenous outgoing and incoming market access terms, or producer and consumer price indices, $\{\Pi_i, P_i\}$ respectively.[14]

We can reformulate the equilibrium equations to solve for the spatial allocation of labor and income, allowing for a rich set of amenity and productivity spillovers. First, we assume welfare equalization, given by $W_j = \frac{w_j}{P_j}u_j$. Then, we allow for localized productivity (A_i) and amenity spillovers (u_i) that depend on the density of workers in a locality. Specifically, productivity is modeled as $A_i = \bar{A}_i L_i^{\alpha}$ and amenities as

[13] The Frechet distribution's properties imply that the inclusive value is given by $p_j = \mathbb{E}_{\varepsilon}\left[\min_{(k,i)\in\mathscr{B}(j)\times N}\frac{t_{kj}\tau_{ik}w_i}{\varepsilon_{ij,k}(\nu)}\right] = \left(\sum_{(k,i)\in\mathscr{B}(j)\times N}\left(t_{kj}\tau_{ik}w_i\right)^{-\theta_j}\right)^{-\frac{1}{\theta_j}} = \left(\sum_{k\in\mathscr{B}(j)} t_{kj}^{-\theta_j}\sum_{i\in N}\left(\tau_{ik}w_i\right)^{-\theta_j}\right)^{-\frac{1}{\theta_j}}$. Note that in the last equation, we can re-arrange the sums to highlight that under a homogenous dispersion parameter ($\theta_j = \theta$), the formula takes on a recursive form and connects expected sourcing prices back to origin i along the network topology, $p_j^{-\theta} = \sum_{k\in\mathscr{N}(j)} t_{kj}^{-\theta}p_k^{-\theta}$ since $p_k^{-\theta} = \sum_{i\in N}\left(\tau_{ik}w_i\right)^{-\theta}$.

[14] Note that $\tilde{\Pi}_k^{-\theta_i} \equiv \left(\sum_j t_{ik}^{\Delta\theta_j}\Pi_i^{\Delta_{i,j}}\left(\gamma_i\right)^{\frac{\theta_j}{\theta_i}-1}\tau_{kj}^{-\theta_j}\frac{\delta_j}{P_j^{-\theta_j}}\right)^{\frac{\theta_i}{\theta_j}}$, where $\tilde{\Pi}_k$ is the approximate price index that controls for how local changes in transport cost, and equilibrium outcomes interact with differences in elasticities of substitution between neighboring nodes and where $\Delta\theta_j \equiv \theta_i - \theta_j$. Notice, that it takes on almost the same form as the producer price index, except for a location i specific coefficient that is applied to each individual element in the sum and depends on the relative magnitude of the substitution elasticities, but equals one for the symmetric case. More information and detailed derivation in the Appendix A.3.

$u_i = \bar{u}_i L_i^{\beta}$, where $\bar{A}_i$ represents the exogenous productivity component at location i, and α measures the sensitivity of productivity to local population L_i (capturing productivity spillovers). Similarly, $\bar{u}_i$ denotes the exogenous utility from residing in location i, while β indicates the extent to which amenities are influenced by the local population (capturing amenity spillovers). This allows us to rewrite the price indices in terms of labor and income allocations (Footnote (1) in Appendix A.2). As a result, we have consumer price index $P_j = \frac{1}{\bar{W}} \bar{u}_j L_j^{\beta-1} Y_j$ and producer price index $\Pi_i = \bar{A}_i L_i^{1+\alpha} Y_i^{-\frac{\theta_i+1}{\theta_i}}$.[15] Substituting both price indices into the equilibrium system in Equation (4), we obtain a system of equation that determines the the spatial distribution of income and labor as well as a global scalar that determines aggregate welfare $\bar{W}$ (see Appendix Section A.3 for further details).

Specifically, given the productivity and amenity fundamentals $\left\{\bar{A}_i, \bar{u}_i\right\}$ as well as the aggregate transport costs, the system of $2N$ equations can be solved for the $2N$ endogenous equilibrium values. The equilibrium system determines the endogenous variables via the interaction of price indices along the network topology, where transport cost is endogenously determined as part of the equilibrium system.

Our framework has three main benefits. First, our approach allows us to nest the choice of different transport modes within the optimal route choice model, with a separate elasticity of substitution across modes from the elasticity across routes. This is possible because our model recasts the equilibrium problem in terms of edge-level outcomes and allows us to incorporate richer edge-level decision making and substitution margins. Second, we can include a rich set of mode-specific congestion patterns, including congestion at intermodal terminals as well as along the primary network. Third, we do not rely on matrix inversion methods and can derive closed-form expressions for equilibrium transport costs, traffic flows, and market access terms, which depends on mode and route choices as well as the topology of the transport network. To focus on multimodal transport networks and incorporate transport mode choice into this framework, we assume a homogeneous route choice elasticity ($\theta_j = \theta$) from now.

3.4 Nested Mode Choice and Aggregate Transport Cost

Previously, we considered a consumer making a route-source decision recursively, where choices are made sequentially along edges of the graph. To expand this framework to incorporate routing decisions across multiple transport modes, we introduce a nested mode-choice structure into the current setup. Now, given the node the consumer has selected, she then chooses the cost-minimizing mode out of all available mode options when travelling to that node (subject to an extreme value distributed cost shock). To fix ideas, consider a consumer in location i who has chosen to route via neighboring node k. At k, she then considers the different modes that are available along edge ik, i.e. $m \in \mathcal{M}_{ik}$, where $\mathcal{M}_{ik}$ is the set of modes available between i and k. One way to represent this is using a multi-layered graph (Figure

[15] We can furthermore rewrite the equations in terms of rescaled variables, $\{y_i, l_i\}$, where we define shares of world income in location i, $y_i \equiv \frac{Y_i}{Y^W}$, and shares of total labor in location i, $l_i \equiv \frac{L_i}{L^W}$. To illustrate this, we did this explicitly for the counterfactual equilibrium system in the Appendix A.8.

5b), where the set of available nodes depends on the existing infrastructure. Each layer in Figure 5b represent a mode-specific network, with the top red layer having fewer nodes than the blue (middle) and yellow (bottom) layers. The consumer selects a mode-specific layer for each edge in the network.

In this setting we assume that the edge-level and mode-specific transport costs for the primary mode is $t_{ik,1}$, but for any secondary mode ($m \neq 1$) a switching cost is imposed, i.e. $t_{ik,m} = s_{ii,m}\iota_{ik,m}s_{kk,m}$ for $m \neq 1$, where $\iota_{ik,m}$ refers to the iceberg transport cost of traversing the edge between node i and k along mode m. Notice that this specification is general and allows for geographies where the secondary modes might connect both primary and secondary nodes. We present the expected aggregate transport cost as the sum of the endogenous mode-specific transport cost, $t_{ik,m}$, that is available between neighboring nodes i and k, i.e. $\mathscr{M}(i,k)$ (see detailed derivations in Appendix A.4):

$$t_{ik} = \mathbb{E}\left[\min_{m\in\mathscr{M}(i,k)}\left\{\frac{t_{ik,m}}{\varepsilon_{ik,m}}\right\}\right] \propto \left(\sum_{m\in\mathscr{M}(i,k)} t_{ik,m}^{-\eta}\right)^{-\frac{1}{\eta}} \tag{5}$$

where the cost-minimizing mode choice is subject to an extreme-value distributed cost shock with dispersion parameter η. The Frechet distribution properties imply the following mode choice probability,

$$\pi_{ik}^{m} = \frac{t_{ik,m}^{-\eta}}{t_{ik}^{-\eta}} \tag{6}$$

This setup allows for a flexible introduction of transport mode choice in a setting where agents are already making endogenous optimal routing and sourcing choices. In our model, mode choice arises as an additional routing decision along a multi-layered graph, where the agent selects the most cost-effective mode for each link. This process leads to an endogenous characterization of the aggregate transport costs, which the agent takes into account when making the overall cost-minimizing routing and sourcing choice. While the introduction of multimodal transport in a segmented transport system adds conceptual complexity, our equilibrium system remains *as tractable* as the one in AA2022. In fact, our model nests AA2022 as a special case when either the dispersion parameter governing mode choice is identical to the one for route choice, $\eta = \theta$, or where only one transport mode is available (see Appendix A.5 for details). The additional step in our framework involves solving for the equilibrium mode-specific traffic.

3.5 Congestion and Traffic Flows

To introduce congestion and derive traffic flows, we proceed in three steps. First, we calculate the aggregate network traffic flows. Second, we determine the mode- and terminal-specific traffic flows. These flows allow us to incorporate congestion into the equilibrium conditions defined in Equation (4). Specifically, we model congestion as switching costs that depend on terminal throughput, creating bottlenecks in the multimodal transport system.

Traffic Flows on the Aggregate Network. We start by constructing the a measure of link intensity—the probability that a route i to j is used and the likelihood of selecting a specific link kl for that route:

$$\pi_{ij}^{kl} = \left(\frac{\tau_{ij}}{\tau_{ik} t_{kl} \tau_{ij}} \right)^{\theta} \tag{7}$$

where in Appendix A.6.1 we show how this formula follows directly from our recursive routing choice framework. We use these derivations to define a gravity equation for aggregate traffic flows on link kl,

$$\Xi_{kl} = t_{kl}^{-\theta} \times P_k^{-\theta} \times \Pi_l^{-\theta} \tag{8}$$

where traffic flows are a function of inward, $P_k^{-\theta}$, and outward market access measures, $\Pi_l^{-\theta}$, as well as the aggregate transport cost for that link across the multimodal transport network.

Traffic Flows on the Mode-Specific Networks. We next characterize the probability of using route ij, the likelihood of selecting link kl as the cost-minimizing option for that route, and the likelihood of selecting mode m as the lowest cost mode for that link. The nested choice implies that this probability is given by, $\pi_{ij,kl,m} = \pi_{ij} \times \pi_{ij}^{kl} \times \pi_{ik}^{m}$ where the π_{ij} is the ij route-sourcing share (Equation (3)), π_{ij}^{kl} is the kl link choice probability conditional on choosing route ij (Equation (7)), and π_{ik}^{m} is the mode share conditional on both route and link choices (Equation (6)). Given this choice probability, we define the mode-specific traffic flows on kl as the gravity equation below (see Appendix Section A.6.2 for details),

$$\Xi_{kl,m} = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \pi_{ij,kl,m} E_j = t_{kl,m}^{-\eta} \times t_{kl}^{\eta-\theta} \times P_k^{-\theta} \times \Pi_l^{-\theta} \tag{9}$$

which connects mode-specific traffic flows to market access measures, as well as mode-specific and link-level transport costs along the link ($t_{kl,m}$ and t_{kl} respectively). Note that the nested formulation implies that mode-specific traffic is the conditional on the mode-specific share of aggregate traffic, i.e. $\Xi_{kl,m} = \pi_{ik}^{m} \times \Xi_{kl}$. Equation (9) is also the mode-specific analogue to Equation (8).

Traffic Flows at Terminals. Here we characterize traffic flows at terminals where switches between the primary and secondary modal network occurs. We derive terminal-level traffic flows in a straightforward manner by summing the mode-specific traffic at any origin terminal k, i.e. $\Xi_{kk,m} = \sum_{l \in \mathscr{F}_m(k)} \Xi_{kl,m} = P_k^{-\theta} \times \sum_{l \in \mathscr{F}_m(k)} t_{kl,m}^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta}$ where $\mathscr{F}_m(k)$ denotes the set of mode-specific successor nodes of k. Combining this with mode-specific transport cost, we obtain the following (Appendix A.6.3 for details),

$$\Xi_{kk,m} = P_k^{-\theta} \times s_{kk,m}^{-\eta} \times \sum_{l \in \mathscr{F}_m(k)} \left(\tau_{kl,m} s_{ll,m} \right)^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta}, \tag{10}$$

This equation differs slightly from the two previous traffic equations due to an additional summation term. This term, a higher-order market access measure, shows how traffic flows through a terminal depends on the sum of traffic generated by nodes that can be accessed via that terminal along the secondary network. This term also captures the centrality of terminals in connecting primary and secondary networks, highlighting their potential to become bottlenecks in the overall transport network.

Congestion on the Primary Transport Network. Having derived traffic flows, we now incorporate congestion on the primary network and then at terminal stations.[16] The congestion model for the primary network is based on AA2022, assuming that travel costs on a link depend on its traffic volume,

$$t_{kl} = \bar{t}_{kl} \left(\Xi_{kl,1} \right)^{\lambda_1}, \tag{11}$$

where λ_1 determines the strength of congestion on the primary network, $\overline{\mathbf{T}} \equiv [\bar{t}_{kl}]$ is the primary network's infrastructure, and $\Xi_{kl,1}$ is the network's link-level traffic flows. As long as $\lambda_1 > 0$, transport cost on a link increases as its traffic flow increases.[17] Using this expression, we first state the primary network's transport costs as a function of its infrastructure, aggregate-level transport costs, and market access terms (first equation below). Next, we combine this transport cost with Equation (8) to express the equilibrium traffic flows (second equation below):

$$t_{kl,1} = \bar{t}_{kl,1}^{\frac{1}{1+\eta\lambda_1}} \times t_{kl}^{\frac{\lambda_1(\eta-\theta)}{1+\eta\lambda_1}} \times P_k^{\frac{-\theta\lambda_1}{1+\eta\lambda_1}} \times \Pi_l^{\frac{-\theta\lambda_1}{1+\eta\lambda_1}}, \quad \Xi_{kl,1} = \bar{t}_{kl,1}^{\frac{-\eta}{1+\eta\lambda_1}} \times t_{kl}^{\frac{\eta-\theta}{1+\eta\lambda_1}} \times P_k^{\frac{-\theta}{1+\eta\lambda_1}} \times \Pi_l^{\frac{-\theta}{1+\eta\lambda_1}} \tag{12}$$

The equilibrium traffic flows on the primary network depend on the fundamental transport capacity and aggregate transport cost of each link ($\bar{t}_{kl,1}$ and t_{kl}), both inward and outward market access terms, as well as the strength of the congestion externality (λ_1). Intuitively, as better market access improves traffic flow on each link, this also increases congestion which mutes its overall impact.

Congestion at Intermodal Terminals. When introducing congestion at terminals, we assume that the direct cost of transiting through a terminal depends on the overall traffic at the terminal,

$$s_{kk,m} = \bar{s}_{kk,m} \left[\Xi_{kk,m} \right]^{\lambda_m}, \quad \text{where} \quad m \neq 1 \tag{13}$$

where λ_m specifies the strength of congestion at terminals, $\overline{\mathbf{S}} \equiv \left[\bar{s}_{kk,m} \right]$ is the switching matrix that connects primary and secondary modes, and $\Xi_{kk,m}$ denotes the traffic at the terminal station. Combining with the expression for terminal traffic (Equation (10)), the transport cost for secondary modes, $t_{kl,m}$, is

$$t_{kl,m} = \bar{s}_{kk,m} \times \tau_{kl,m} \times \bar{s}_{ll,m} \times P_k^{-\theta\lambda_m} \times \Pi_l^{-\theta\lambda_m} \times P_{k,m}^{-\theta\lambda_m} \times \Pi_{l,m}^{-\theta\lambda_m}, \quad \text{where} \quad m \neq 1 \tag{14}$$

where $P_{l,m} \equiv \sum_{k \in \mathscr{B}(l)} t_{kl,m}^{-\eta} \times t_{kl}^{\eta-\theta} \times P_k^{-\theta}$ and $\Pi_{k,m} \equiv \sum_{l \in \mathscr{F}(k)} t_{kl,m}^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta}$ define mode-specific market access measures. Intuitively, if $\lambda_m > 0$, transport cost increases with increasing traffic at the terminal. By combining Equation (14) with our modal traffic expression (Equation (9)), we determine the traffic flows for the secondary modes while taking into account endogenous congestion at terminals,

$$\Xi_{kl,m} = \bar{s}_{kk,m}^{-\eta} \times \tau_{kl,m}^{-\eta} \times \bar{s}_{ll,m}^{-\eta} \times P_{k,m}^{\eta\theta\lambda_m} \times \Pi_{l,m}^{\eta\theta\lambda_m} \times t_{kl}^{\eta-\theta} \times P_k^{-\theta(1-\eta\lambda_m)} \times \Pi_l^{-\theta(1-\eta\lambda_m)}, \quad \text{where} \quad m \neq 1 \tag{15}$$

While increases in both incoming and outgoing market access measures generally increases traffic

[16] While it is theoretically feasible to extend the framework to include congestion on the secondary network, we have opted to model congestion at terminals. This choice reflects the fact that bottlenecks then to occur at the intermodal transfer with a prime example being port congestion. Additionally, incorporating congestion at the route-level for secondary networks would come at the cost of added complexity and decreased traceability.

[17] While somewhat different to the more commonly used Bureau of Public Roads (BPR) function (Boyles, Lownes and Unnikrishnan, 2021), it is both analytically convenient and can be micro-founded in a simple model where transport costs are log-linear in travel time and speed is a log-linear function of traffic congestion as shown in AA2022.

flows (Equation (8)), we highlight how these market access terms contribute to traffic flows on both the primary and secondary networks while taking congestion into account (Equations (12) and (15) respectively). Specifically, terminal congestion (λ_m) can reduce the appeal of multimodal routes and consequently decrease traffic on the secondary network (Equation (15)). This highlights the significant role terminals can have as potential bottlenecks within a multimodal network. In Section 4.2, we estimate the terminal congestion elasticity by focusing on one important type of intermodal terminal: ports.

3.6 Substitution and Complementarity within the Multimodal Network

We now demonstrate the substitution and complementarity effects within our model by examining how a mode-specific cost change affects the equilibrium traffic flows of other modes. The net effect of a transport cost change for mode m' on segment kl $t_{kl,m'}$ on the traffic flows of an alternative mode m on the same segment ($\Xi_{kl,m}$, Equation (9)) can be decomposed into two components: (1) a direct substitution effect on alternative mode m's traffic on the same link, and (2) an indirect complementarity effect that impacts the consumer and producer price indices across all modes (Appendix A.7 for details),

$$\frac{d \ln \Xi_{kl,m}}{d \ln t_{kl,m'}} \approx \left[\underbrace{\eta}_{\substack{\text{Direct Substitution} \\ \text{Effect}}} \quad \underbrace{-\theta \left(1 + \left[\tilde{\Omega}^{t,P}\right]_{kl} + \left[\tilde{\Omega}^{t,\Pi}\right]_{kl}\right)}_{\substack{\text{Indirect Complementarity} \\ \text{Effect}}} \right] \omega_{kl,m'} \tag{16}$$

where η is the elasticity of modal substitution, θ is the elasticity of route and trade substitution, $\omega_{kl,m'} \equiv \frac{\Xi_{kl,m'}}{\sum \Xi_{kl,m'}}$ is the traffic share for mode m' on link kl, while $\tilde{\Omega}^{t,P} \equiv \left(\mathbf{I} - \Omega^{t,P}\right)^{-1} \Omega^{t,P}$ and $\tilde{\Omega}^{t,\Pi} \equiv \left(\mathbf{I} - \Omega^{t,\Pi}\right)^{-1} \Omega^{t,\Pi}$ are reduced-form general equilibrium elasticities, which depend on square matrices that determine how transport costs affect consumer and producer price indices across the transport network, i.e. $\Omega^{t,P} = \left[\frac{t_{kl}^{-\theta} P_k^{-\theta}}{P_l^{-\theta}}\right]_{kl}$ and $\Omega^{t,\Pi} = \left[\frac{t_{kl}^{-\theta} \Pi_l^{-\theta}}{\Pi_k^{-\theta}}\right]_{kl}$ respectively. These terms, taking on the form of Leontief inverses, measure the centrality of segments across the transport network.

The first term in Equation (16) describes the effects of direct modal substitution. Consider a decrease in the cost of mode m' on segment kl. Given that agents make nested cost-minimizing choices across transport modes, they will substitute towards the now cheaper mode m' and away from alternative mode m. The extent of this substitution is governed by the elasticity of modal substitution η and is proportional to the traffic share of the alternative mode. Referring back to our example from the introduction, an improvement in the highway network reduces road transport costs, making rail transport relatively more expensive and resulting in a shift in traffic from rail to road.

The second term in Equation (16) describes the indirect complementarity effect. A decrease in the cost of mode m' on segment kl lowers the aggregate transport cost for that segment, benefiting routes and locations that are likely to use it for transport and trade. This leads to two outcomes: (1) increased aggregate traffic for segment kl, and (2) additional effects from improvements in consumer and produce

price indices—inward and outward market access respectively.[18] This latter effect reflects the broader general equilibrium adjustments in the economy. Since price indices are recursively related, transport cost improvements spatially propagate across the network. The strength of this effect depends on the elasticity of route and trade substitution, θ, the centrality of the segment, and is, like the substitution effect, proportional to the traffic share of the alternative mode. The overall traffic impact then depends on the balance of the two opposing effects: the direct substitution effect dominates when η is larger than both θ and the general equilibrium elasticities ($\tilde{\Omega}^{t,P}$ and $\tilde{\Omega}^{t,\Pi}$), while the indirect complementarity effect dominates for segments that are highly central to the network. In our earlier example, better road access indirectly improves market access of all locations, thereby increasing transport demand across all modes, including rail. We discuss these insights in light of our estimation strategy for η in Section 4.1.

This decomposition is not just interesting for its economic insights, but provides guidance for our applications. First, it illustrates our notion of bottlenecks. Segments that have fewer alternatives and limited capacities can limit the overall efficiency of the transport network. Through the lens of our model those are locations with sizeable complementarity effects across the economy. Second, it illustrates environmental implications of changes within the network. Since transport modes differ in their emission levels—trucks, for example, emit eight times more CO_2 per ton-mile than rail (CBO, 2022)—our framework can uniquely evaluate the environmental outcomes associated with the substitution and complementarity effects resulting from mode-specific investments or disruptions.

4 Estimation: Modal Substitution and Terminal Congestion

In order to quantify our model, we require two key elasticities. First, to motivate the mode choice channels in our model, we require an elasticity of substitution between transport modes (η). In Section 4.1 we estimate this elasticity by examining how traffic shifts between modes in response to cost changes in a one mode. Second, the counterfactual equilibrium in our model depends on the strength of the congestion forces within the network, particularly at intermodal terminals (λ_m).[19] In Section 4.2, we estimate λ_m by analyzing the relationship between ship dwell times and port traffic. We demonstrate that this elasticity, though estimated for port terminals, also applies to inland intermodal rail terminals.

4.1 Estimation of Elasticity of Modal Substitution

We start by motivating our approach to estimate the elasticity of modal substitution η, and its structural interpretation through the lens of our model, as discussed in Section 3.6. Next, we explain how

[18] We can derive closed-form expressions for these market access terms with Equation (4), as detailed in Appendix A.7. To keep things simple, we derive this effect in the absence of congestion. In principle, we can also allow for congestion. This would add additional terms into the implicitly differentiated equilibrium conditions.

[19] There are two congestion forces in our model: on the primary network and at intermodal terminals (λ_1, λ_m). We rely on prior literature for λ_1 and focus on estimating λ_m in this paper.

we estimate this elasticity by adopting the research design from Duranton and Turner (2011) and incorporating the focus on relative modal responses from Hummels and Schaur (2013) and Lugovskyy, Skiba and Terner (2022). Third, we conduct several robustness checks to validate our results.

First, consider a mode-specific transport cost decrease, like a local road improvement. Following our discussion in Subsection 3.6, this improvement has both direct and indirect effects. Directly, lower road costs will substitute traffic from rail or barges to road. The strength of this substitution effect directly depends on our parameter of interest η. Indirectly, lower road costs will reduce aggregate trade costs and improve overall market access across all modes (the complementarity effect). Since we are interested in estimating η, the complementarity effect constitutes confounding variation, as changes in mode-specific traffic are endogenous to changes in both incoming and outgoing market access (Equation (16)). We address this, and demonstrate more formally in the Appendix B.2, by examining the instrumented changes in relative modal traffic responses (modal split) to mode-specific cost changes rather than the mode-specific traffic responses directly (Equation (29), Appendix B.2).

Empirical Strategy. Our approach builds on Duranton and Turner (2011) (hereby abbreviated to DT2011), who found that improving road infrastructure by increasing lane kilometers of interstate highways, increases the vehicle-kilometers traveled (VKT) in US metropolitan cities particularly for commercial freight trucks. To address the endogeneity between VKT demand and road stock changes at the metropolitan statistical areas (MSAs) level, they utilize three instruments: kilometers of preliminary interstate highway from the 1947 highway plan in each MSA (Baum-Snow, 2007; Michaels, 2008), kilometers of 1898 railroads in each MSA, and 1528-1850 exploration routes.[20]

By aggregating our waybill rail data to the MSA level and matching it with their data, we estimate the elasticity of modal substitution by studying how traffic shifts between modes in response to cost changes in a one mode, such as the effect of reduced road costs on relative rail and road traffic flows.[21] We start with the following OLS specification with fixed effects from DT2011:

$$\ln Y_{cy} = \alpha \ln \text{Interstate Highway Lanes}_{cy} + \psi C_{cy} + \zeta_c + \iota_y + \varepsilon_{cy} \tag{17}$$

where $\ln Y_{cy}$ is log traffic use outcomes for MSA c in year y, $\ln \text{Interstate Highway Lanes}_{cy}$ is log number of interstate highway lanes going through MSA c which proxies for the road infrastructure for that MSA in year y. C_{jt} are DT2011 city-specific time-varying controls including population, geography, census divisions, and socioeconomic characteristics. ζ_c is MSA-level fixed effects, and ι_y is year fixed effects.

We measure road traffic use with truck VKT from DT2011, due to our focus on commercial freight transport. To match truck VKT which is measured in vehicle-kilometers traveled, we measure rail traffic

[20] Using the same three instruments, Duranton and Turner (2012) finds that a 10% increase in interstate highways increases city employment by about 1.5% between 1983 and 2003. With the same instruments, Duranton, Morrow and Turner (2014) finds that more highways lead to cities specializing in exporting goods with higher weight-to-value ratios.

[21] After the matching process, we have 221 MSAs which is 7 less compared to DT2011 (see Section B.1.5 for more information). Despite the slight difference in observations, we replicate the DT2011 results (Table A.2): our OLS and IV estimates have the same sign and are within one standard deviation of their results (Columns (6) to (10), Table 9, DT2011).

use in total railcar-kilometers-traveled for each city. Our detailed rail data also allows us to calculate the rail traffic use by destination—incoming rail traffic use—and by origin—outgoing rail traffic use. We also observe rail car weight, and can construct an alternative measure of rail traffic use by weight in weight-kilometers-traveled. In order to measure relative changes between truck and rail use for US cities in response to road infrastructure improvements, we take the ratio of rail traffic use (measured in railcar-km or weight-km) to road traffic (measured in truck vehicle-km) for city c in year t $\left(\frac{\text{Rail Traffic Use}_{cy}}{\text{Road Traffic Use}_{cy}}\right)$.

OLS Results We start by describing the mode-specific results for road and rail, and then the relative modal split responses. As mentioned earlier, a decrease in road cost will increase its traffic flows (Equation (8)). With our matched rail and truck data, we first show that we can replicate the significant positive road traffic use estimates from DT2011. Our OLS estimates (Columns (1) and (2), Table A.2) are within one standard error of their results (Columns (6) and (7), Table 9, DT2011). Second, using our rail traffic use measure, we find a noisy negative relationship with the decrease in road costs (Columns (1) and (2), Table A.4). These OLS coefficients are imprecisely estimated due to the two opposing substitution and complementarity forces shown in Equation (16) and discussed in Subsection 3.6: while a decrease in road costs raises rail cost relative to road, *reducing* rail traffic, it can also lower aggregate transport cost for the city and improve its overall market access, *increasing* demand for rail traffic.

Next, we study the relative modal split responses in rail and road traffic use. In Columns (1) and (2) of Table 1, we find that relative rail to truck traffic use has a negative and significant relationship with road infrastructure improvement in US cities. This result is robust to including city-level and year-level fixed effects, as well as time-varying city-level controls like socioeconomic characteristics and population. Note that this result, a coefficient of -1.4, is trivially explained by the inverse of road traffic use estimate: this estimate is lower in absolute terms from road traffic use OLS estimate of +1.6 from Table A.2 (Columns (1) and (2)). This negative relationship is robust to measuring rail traffic use in terms of rail weight as well (Columns (1) and (2), Table A.5).

IV strategy. Since transport infrastructure and traffic use may be simultaneously determined, we required an instrumental variable approach in order to identify the causal effects of road infrastructure on rail traffic use and modal substitution. We employ the three instruments by DT2010 to predict the stock of roads in US cities and estimate the following two-stage least-squares IV regression:

$$\begin{aligned} \ln \text{Interstate Highway Lanes}_{cy} &= \eta_1 \ln \text{Instruments}_c + \kappa C_{cy} + \iota_y + \nu_{cy} \\ \ln Y_{cy} &= \eta_2 \ln \text{Interstate Highway Lanes}_{cy} + \phi C_{cy} + \iota_y + \mu_{cy} \end{aligned} \tag{18}$$

where $\ln Y_{cy}$ is log traffic use outcomes for MSA c in year y, $\ln \text{Instruments}_c$ is the set of three instruments discussed earlier, $\ln \text{Interstate Highway Lanes}_{cy}$ is log number of interstate highway lanes going through each city c which proxies for the road infrastructure for that city in year y. C_{jt} are city-specific time-varying controls including population, physical geography, census divisions, and socioeconomic characteristics from DT2011, and ι_y is year fixed effects.

Table. 1. Elasticity of Modal Substitution

	(1)	(2)	(3)	(4)	(5)
Rail to Road Traffic Use	OLS	OLS	IV	IV	IV
Interstate Highway Lane KM	-1.432	-1.432	-0.867	-1.249	-1.099
	(0.195)	(0.196)	(0.376)	(0.388)	(0.364)
Population		-0.150	0.699	1.092	0.891
		(0.337)	(0.289)	(0.328)	(0.306)
Geography				✓	✓
Census Divisions				✓	✓
Socioeconomic Characteristics		✓			✓
MSA FE	✓	✓			
Year FE	✓	✓	✓	✓	✓
Observations	658	658	658	658	658
R-squared	0.88	0.88	-0.03	0.23	0.27
KP F-stat			14.48	10.76	10.04

Notes: Robust standard errors clustered by MSAs in parentheses. All variables are in logs. Rail traffic use, measured in railcar-kilometers, is constructed using confidential rail waybill data. Truck traffic use (in vehicle-kilometers) and other variables are from Duranton and Turner (2011). Instruments are 1835 exploration routes, 1898 railroad route kilometers, and 1947 planned interstate highways. 658 observations corresponding to 221 MSAs for each regression. See Table A.3 for first stage regressions.

The validity of the IV strategy requires that the instrument be uncorrelated with unobserved changes in both road and rail traffic use, including changes in the relative mode-specific market access (Appendix B.2), conditional on the control variables and fixed effects in Equation (18). As detailed in DT2011 and later papers, the first instrument, the 1947 highway plan, aimed to connect major 1940s population centers rather than anticipate future traffic or population growth. The second instrument, the 1898 railroad network, was developed by for-profit private companies operating during that time in a smaller agricultural economy, rather than to anticipate demand for transport 100 years later. The third instrument, the 1528-1850 exploration routes, were expeditions to find efficient paths between places, instead of anticipating transport demand at cities 150 years later. Additionally, controlling for historical population and physical geography helps address potential confounders in these instruments.

Results: Elasticity of Modal Substitution. We first show that we can replicate the road traffic use IV results in DT2011 with our matched data set: our IV estimates in Columns (3) and (4) (Table A.2) are within one standard error of their results (Columns (9) and (10), Table 9, DT2011). Similar to them, we find that the IV estimates are slightly higher than the OLS estimates. Our results are also robust to including socioeconomic characteristics (Column (5), Table A.2). Next, we find that road infrastructure has a positive and noisy impact on rail traffic (Columns (3) to (5), Table A.4), suggesting that the complementarity effect dominates the substitution effect (Equation (16), Subsection 3.6).

Turning to the modal split responses, we find that road infrastructure improvements result in a larger increase in road traffic use relative to the rail increase—resulting in a decrease in the ratio of rail to road traffic use. A 1 percent increase in interstate highways causes a 0.9-1.2 percent decrease in rail to road traffic use (Columns (3) to (5), Table 1). We use the estimate from the most conservative specification as

our elasticity of modal substitution, $\eta = 1.099$ (Column (5), Table 1). Again, this result is not trivially driven by the inverse of the road traffic use estimate: our estimates are lower in absolute terms and beyond one standard error from the road traffic use estimates, where a 1 percent increase in interstate highways causes a 1.7-2.1 percent increase in road traffic use (Columns (3) to (5), Table A.2).

We compare our elasticity of modal substitution estimate to existing literature here. Most of the economics literature focus on the substitution between ocean and air freight (Harrigan, 2010; Hummels and Schaur, 2013; Lugovskyy, Skiba and Terner, 2022) or bulk and container shipping within ocean freight (Coşar and Demir, 2018).[22] Compared to studies on ocean and air freight, our estimates are generally lower, likely because the high costs of air shipping leads to a higher elasticity of substitution.[23] Studies in the transportation literature have estimated modal substitution elasticities between rail and truck, and these estimates fall within the confidence interval of our baseline estimate (Oum, 1979, 1989; Beuthe, Jourquin and Urbain, 2014).[24]

Robustness Checks. We conduct several robustness checks on our elasticity of modal substitution. First, using an alternative measure of rail traffic by weight, we find estimates that are within one standard error of our baseline (Columns (3) to (5), Table A.5). Second, since our rail traffic data is available at a more disaggregated level, we can examine incoming and outgoing rail traffic use separately. Comparing the ratio of directional rail traffic use to road traffic use, we find that these directional results are within one standard error of our baseline results, suggesting that both incoming and outgoing relative rail traffic use respond similarly to road infrastructure improvements (Columns (2) and (3) for incoming and outgoing traffic respectively, Table A.6). Third, combining both the alternative weight measure with directional rail traffic use, we find similar results within one standard error of our baseline (Columns (4) and (5) for incoming and outgoing traffic respectively, Table A.6).

Fourth, while many of the 1898 railroads were abandoned and turned into roads (Duranton and Turner, 2011, 2012; Duranton, Morrow and Turner, 2014), we further test the robustness of our results by removing the 1898 railroads as an instrument. Both elasticities using our carload and weight-based measures retain the same sign and are within one confidence interval of our baseline results (Columns (6) and (7) for carload- and weight-based rail traffic use respectively, Table A.6).

[22] Instead of estimating an elasticity of modal substitution, Coşar and Demir (2018) estimates the cost structures of container and breakbulk shipping to quantify the impact of containerization in trade volumes.

[23] Hummels and Schaur (2013) estimates substitution elasticities of 2.7-6.5 for air versus ocean shipping, while Lugovskyy, Skiba and Terner (2022) reports an IV estimate of 10.3 for air versus containerized ocean shipping. Focusing on unit values, Harrigan (2010) finds that a 1% increase in value per weight raises the probability of air shipment by 0.2 percentage points.

[24] Differences in empirical strategies and context complicate direct comparisons. Beuthe, Jourquin and Urbain (2014) finds a rail ton-mile elasticity of 0.59 with respect to road cost reductions. Using 1970 Canadian freight flow data and a translog demand system, Oum (1979) estimates an rail-truck elasticity of 1.4-1.57 for eight selected commodities and Oum (1989) estimates an elasticity of 1.19 for all goods.

4.2 Estimation of Intermodal Terminal Congestion

In this subsection, we first describe our approach to estimate the congestion elasticity λ_m at intermodal terminals, using minute-level AIS data to measure ship loading and unloading times (dwell times) and examining how these dwell times respond to port traffic. To account for the potential endogeneity of ship dwell times, we employ a shift-share instrument. Next, we describe our results and compare our estimates to the literature. Lastly, we conduct robustness checks to test the validity of our results.

Note that we apply λ_m to all intermodal terminals in our model, including inland rail terminals, even though it is based on ports. This is because the highly detailed AIS panel data, which enables us to precisely capture ship dwell times and include extensive fixed effects in our estimation, is not available at the same granularity for rail terminals. Despite this, using a smaller railcar dwell time dataset, we find a significant positive link between railcar dwell times and port traffic, highlighting the interconnectedness of transport modes and suggesting that port congestion can proxy for congestion at other terminals.

Empirical Strategy. We measure port congestion by estimating the elasticity of ship dwell time with respect to port traffic. We estimate the following regression:

$$\ln \text{Ship Dwell Time}_{spdmy} = \beta_1 \ln \text{Port Traffic}_{pdmy} + \varphi_{sp} + \rho_{py} + \psi_{dmy} + \varepsilon_{spdmy} \tag{19}$$

where $\text{Ship Dwell Time}_{spdmy}$ is the number of hours ship s spent at port p on day of the week d month m and year y, $\text{Port Traffic}_{pdmy}$ is the 28-day moving average amount of port traffic at port p ending on day d month m and year y, φ_{spy} is ship-port fixed effects, ρ_{py} is port-year fixed effects and ψ_{dmy} is day-month-year fixed effects. The key parameter of interest, β_1, captures the elasticity of ship dwell times with respect to port traffic. Standard errors are two-way clustered at the ship and port level.

The ship-port fixed effects control for fixed characteristics at the ship-level, port-level, and ship-port-level. Fixed ship characteristics like ship sizes are important to control for since larger ships have large capacities and naturally require more time to load and unload. Fixed port characteristics like the geographical proximity of ports to large trading partners or to multiple major cities domestically can also play a role in attracting more trade which can contribute to higher loading and unloading times (as seen in Figure 4b for LA and Long Beach). Fixed characteristics at the ship-port level include time-invariant comparative advantage differences for different ports that result in larger ships being received at these ports which increases dwell times, for example ports with deep natural harbors. The port-year fixed effects control for time-varying port characteristics which includes potential technology or infrastructure changes over time that ports can undertake that might affect ship dwell times, for example technology upgrades at ports over time to accommodate larger ships. Additionally, the day-month-year fixed effects control for aggregate events that impacts all ships.

OLS Results We find that a one percent increase in port traffic is correlated with a statistically significant increase in ship dwell times by 0.1 percent (Column (2), Table 2). This elasticity is within one standard

error of the baseline when estimated with ship fixed effects separately (Column (1), Table 2). West Coast ports, with their history of strikes, deeper harbors, and large volumes of US-Asia trade (via LA and Long Beach), often experience longer ship dwell times. We show that the OLS coefficient for West Coast ports is indeed more than two times larger (Column (1), Table A.7). Including a pandemic period indicator (post-March 2020) to estimate separate coefficients during this time period, we find that the pre-March 2020 estimate is within one standard error of the baseline (Column (2), Table A.7), while the post-March 2020 estimates is slightly higher but still within one confidence interval of the baseline. Using shorter periods of the moving average calculation, we find that the ship dwell times are still statistically significantly correlated with port traffic but the magnitude of the coefficients are smaller (Table A.8).

IV Strategy. In order to establish the causal impact of port traffic on ship dwell times, we require demand shifter for port traffic that is uncorrelated with unobserved ship dwell times determinants ε_{spdmy}. Our IV strategy uses aggregate compositional changes for imports at the region- and product-level into the top 30 US ports to predict demand for port services at each port:

$$\text{Port Trade Exposure}_{pmy} \equiv \sum_{O}\sum_{N} X_{on\backslash p,my} \times \omega_{onp,2003} \tag{20}$$

where shift-share IV is the weighted sum of top 30 US ports imports from origin o and product n excluding port p at month m and year y. The first term in Equation (20) is the contemporaneous change in overall demand for top US ports while leaving out its own port p, while the second term is the lagged region- and product-level share by at least 13 years from the start of the available sample period (year 2003). The weights sum to one across products and origins for each port-year observation. The Census Bureau monthly trade data allows us to match this IV with our 28-day moving average of port traffic.

For this IV strategy to be valid, the port trade exposure measure for each port has to be generally uncorrelated with unobserved ship dwell times determinants for the ship servicing these ports after controlling for fixed characteristics at the ship-port level, time varying characteristics at the port-level, and aggregate time-varying events that impact all ships. Additionally, we establish that our IV strategy has a strong first stage: the port trade exposure measure has a statistically significant and positive relationship with port traffic (Column (5) and (6), Table 2).

Results: Elasticity of Intermodal Terminal Congestion. We find that a one percent increase in port traffic results in a significant 0.24 percent increase in ship dwell time (Column (6), Table 2). This result is within one standard error of the specification with ship fixed effects separately (Column (4), Table 2). This congestion elasticity forms the basis for congestion elasticity at terminals, parameter λ_m (Equation (13)). Since this elasticity is measured in time while Equation (13) is expressed in cost terms, we convert this time elasticity into a cost measure using the ad valorem tariff estimate for an additional day in transit from Hummels and Schaur (2013), resulting in λ_m=0.096.

Without directly comparable elasticities on intermodal terminal congestion, we compare our estimate

Table. 2. Congestion Elasticity of Port Traffic with respect to Ship Dwell Times

	(1) OLS	(2) OLS	(3) First-Stage	(4) IV	(5) First-Stage	(6) IV
Port Traffic	0.093	0.100		0.264		0.236
	(0.011)	(0.010)		(0.095)		(0.091)
Port Trade Exposure			0.229		0.228	
			(0.013)		(0.014)	
Day-Month-Year FE	✓	✓	✓	✓	✓	✓
Port-Year FE	✓	✓	✓	✓	✓	✓
Ship-Port FE		✓			✓	✓
Ship FE	✓		✓	✓		
Observations	90516	90516	90516	90516	90516	90516
First Stage KP-F				299.03		274.05

Notes: Robust standard errors in parentheses are two-way clustered at the ship and port level. All variables are in logs. Port traffic is the 28-day moving average of total daily net tonnage at the port. Weighted by ship net tonnage.

with two strands of literature. First is the literature on processing times at ports. Using detailed Peruvian data, Carballo et al. (2021) estimates an import processing cost elasticity of 0.06 which is within the confidence interval of our baseline estimate.[25] Second, we compare our estimates to congestion elasticities from the urban literature which focuses on road congestion in cities. Existing US estimates range from 0.03-0.11 which also fall within one confidence interval of our baseline estimate (Couture, Duranton and Turner, 2018; Allen and Arkolakis, 2022; Akbar et al., 2023; Duranton and Puga, 2023).[26]

Robustness Checks. As noted earlier, West Coast ports typically have longer dwell times due to their location and history of strikes. Indeed, we find that congestion elasticity is higher for these ports (Column (3), Table A.7), but it is still within one confidence interval of our baseline result. As additional checks, we examine the before and after the pandemic period separately. While the post-pandemic period has a higher congestion elasticity—as we would expect, both elasticities are within one standard error of the baseline (Columns (4) and (5) respectively, Table A.7).

Although our congestion elasticity λ_m is applied to all intermodal terminals within the multimodal network, it is estimated solely from AIS vessel traffic data at ports. This is due to the exceptional level of detail provided by this panel data, which allows us to capture ship dwell times with precision and incorporate an extensive range of fixed effects. To the best of our knowledge, comparable rail station dwell times data with the same level of granularity is not available. However, by using a smaller sample

[25] Carballo et al. (2021) estimates that a 1% increase in median import processing time reduces import values by 0.24%, giving a processing cost elasticity of 0.06 with a demand elasticity of 4. With export processing data on 98 countries, Djankov, Freund and Pham (2010) finds that a 1% increase in delivery time from factory to port reduces trade by 0.4%.

[26] We focus on US estimates for comparison purposes. Couture, Duranton and Turner (2018) estimates a congestion elasticity of 0.11 associated with US metro area population while Duranton and Puga (2023) estimates an elasticity of 0.04 using residualized city speed for 180 US metropo areas. Akbar et al. (2023), with a global dataset on vehicle travel speeds, estimate a congestion elasticity with respective to income of 0.032 and provides an in-depth discussion of on existing congestion elasticity estimates. AA2022 estimate a road congestion elasticity of 0.092.

of rail dwell times matched to port traffic data for nearby rail stations, we find a statistically significant and positive relationship between port traffic and rail dwell times at these rail stations (Table A.9, Appendix Section B.5).[27] This finding highlights the interconnectedness of transport modes within the multimodal network, where congestion at ports likely leads to delays at other key intermodal terminals. Consequently, port congestion can serve as a proxy for understanding congestion impacts at intermodal terminals across the broader multimodal network.

5 Calibration and Validation

To apply our framework for evaluating the overall impact of infrastructure investments and disruptions within the multimodal transport network, we begin by presenting two propositions that in tandem define a set of equations to compute counterfactuals for arbitrary changes of the multimodal transport network. Next, we discuss how we construct the US multimodal transport network using detailed traffic and geography data, and the calibration of key parameters in the model. We then turn towards providing a validity check of our framework by comparing the observed value of bilateral trade flows between cities, at the transport mode-level, to our model predicted trade flows. Lastly, we conduct counterfactuals on highway segment improvements to compare the results within our multimodal transport framework to a unimodal one per existing literature.

5.1 Counterfactual Equilibrium

We now derive a system of equations that determines the counterfactual equilibrium for arbitrary mode-specific cost changes within the multimodal network. This allows us to evaluate the welfare impact of investments or disruptions in a setting where agents make sophisticated routing and mode choices while also allowing for a rich characterization of congestion across the multimodal transport system. To do so, we first extend the equilibrium system in Equation (4) to allow for mode choice as in Section 3.4 and for congestion as in Section 3.5. Following Dekle, Eaton and Kortum (2008), we then employ 'Hat Algebra' to express the equilibrium in terms of changes of the endogenous variables. We denote the changes in variables with hats, $\hat{\gamma}_i \equiv \frac{\gamma_i'}{\gamma_i}$. The structure of this model allows for a convenient separation of the edge-specific transport equilibrium *given* market access terms, and the determination of the aggregate equilibrium *given* (aggregate) transport cost.

We present the counterfactual equilibrium in terms of two different parts: The first one describing the determination of the aggregate equilibrium and the second describing a subroutine that determines the transport equilibrium at the edge-level. Detailed derivations for both are provided in Appendix A.8.

[27] See Appendix Section B.1.6 for further details on the rail station dwell times data.

Proposition 1 *Consider an economy in equilibrium with a primary transport network,* $\overline{\mathbf{T}} \equiv \left[\bar{t}_{kl,1}\right]$*, secondary transport network (m),* $\overline{\mathbf{T}}_m \equiv \left[\tau_{kl,m}\right]$*, and intermodal terminal network connecting both primary and secondary networks (m),* $\overline{\mathbf{S}}_m \equiv \left[\bar{s}_{kk,m}\right]$*. Consider any changes either in the primary infrastructure network,* $\hat{\bar{t}}_{kl,1}$*, in the intermodal terminal network,* $\hat{\bar{s}}_{kk,m}$*, or in the secondary transport network,* $\hat{\tau}_{kk,m}$*. Given observed aggregate and mode-specific traffic flows* $\left(\Xi_{ik}, \Xi_{ik,m}\right)$*, economic activity in the geography* $\left(Y_i, E_j\right)$*, and parameters* $\{\alpha, \beta, \theta, \lambda_1, \lambda_m, \eta\}$*, the equilibrium change in economic outcomes* $\left(\hat{y}_i, \hat{l}_i, \hat{\chi}\right)$ *is the solution to the following :*

1. ***Aggregate Equilibrium.*** *Changes in incoming and outgoing market access terms* $\left\{\hat{\Pi}_i, \hat{P}_j\right\}$ *given by:*

$$\begin{aligned}
\hat{P}_j^{-\theta} &= \left(\frac{\gamma_j}{\gamma_j + \sum_{k \in \mathscr{B}(j)} \Xi_{kj}}\right) \frac{\hat{\gamma}_j}{\hat{\Pi}_i^{-\theta}} + \sum_{k \in \mathscr{B}(j)} \left(\frac{\Xi_{kj}}{\gamma_i + \sum_{k \in \mathscr{B}(j)} \Xi_{kj}}\right) \hat{t}_{kj}^{-\theta} \hat{P}_k^{-\theta} \\
\hat{\Pi}_i^{-\theta} &= \left(\frac{\delta_i}{\delta_i + \sum_{k \in \mathscr{F}(i)} \Xi_{ik}}\right) \frac{\hat{\delta}_i}{\hat{P}_i^{-\theta}} + \sum_{k \in \mathscr{F}(i)} \left(\frac{\Xi_{ik}}{\delta_i + \sum_{k \in \mathscr{F}(i)} \Xi_{ik}}\right) \hat{t}_{ik}^{-\theta} \hat{\Pi}_k^{-\theta}
\end{aligned} \tag{21}$$

   where $\hat{t}_{ki}^{-\theta} = F\left(\left\{P_j, \Pi_i\right\}\right)$ *is the change in endogenous transport costs derived from the transport equilibrium in Proposition (1.2),* $\delta_i = \gamma_i = y_i$ *represents the share of income in location i, and changes in the market access terms* $\hat{P}_j$ *and* $\hat{\Pi}_i$ *are equal to changes in income and labor densities,* $\hat{P}_j = \hat{y}_j \hat{l}_j^{\beta-1} \hat{W}^{-1}$ *and* $\hat{\Pi}_i = \hat{l}_i^{\alpha+1} \hat{y}_i^{-\frac{\theta+1}{\theta}}$ *where* y_i *is the share of world income in i,* $y_i \equiv \frac{Y_i}{Y^W}$*, and* l_i *is the share of total labor in i,* $l_i \equiv \frac{L_i}{L^W}$*.*

2. ***Transport Equilibrium.*** *Given* $\left\{\hat{P}_i, \hat{\Pi}_i, \hat{\gamma}, \hat{\delta}\right\}$*, the equilibrium change in endogenous aggregate transport costs* $\left(\hat{t}_{ik}^{-\theta}\right)$ *is the solution to the edge-level transport equilibrium and given by:*

$$\hat{t}_{ik}^{-\theta} = \left(\sum_{m \in \mathscr{M}(i,k)} \frac{\Xi_{ik,m}}{\Xi_{ik}} \hat{t}_{ik,m}^{-\eta}\right)^{\frac{\theta}{\eta}} \tag{22}$$

   and changes in the endogenous transport cost on the primary mode and any alternative mode are given by, $\hat{t}_{kl,1} = \hat{\bar{t}}_{kl,0}^{\frac{1}{1 \cdot \lambda_1}} \hat{t}_{kl}^{\frac{\lambda_1(\eta-\theta)}{1+\eta\lambda_1}} \hat{P}_k^{\frac{-\theta\lambda_1}{1+\gamma\lambda_1}} \hat{\Pi}_l^{\frac{-\theta\lambda_1}{1 \cdot r\lambda_1}}$ *and* $\hat{t}_{kl,m} = \hat{\bar{s}}_{kk,m} \hat{\tau}_{kl,m} \hat{\bar{s}}_{ll,m} \hat{P}_k^{-\theta\lambda_m} \hat{\Pi}_{k,m}^{-\theta\lambda_m} \hat{\Pi}_l^{-\theta\lambda_m} \hat{P}_{l,m}^{-\theta\lambda_m}$ *where the mode specific market access terms are defined as,* $\hat{P}_{l,m}^{-\theta} \equiv \sum_{k \in B(l)} \frac{\Xi_{kl,m}}{\sum_{k \in \mathscr{B}(l)} \Xi_{kl,m}} \hat{t}_{kl,m}^{-\eta} \hat{t}_{kl}^{\eta-\theta} \hat{P}_k^{-\theta}$ *and* $\hat{\Pi}_{k,m}^{-\theta} \equiv \sum_{l \in \mathscr{F}(k)} \frac{\Xi_{kl,m}}{\sum_{l \in \mathscr{F}(k)} \Xi_{kl,m}} \hat{t}_{kl,m}^{-\eta} \hat{t}_{kl}^{\eta-\theta} \hat{\Pi}_l^{-\theta}$*.*

Proposition 1 constructs the aggregate counterfactual spatial equilibrium solving for market access terms while simultaneously solving for the modal traffic assignment. The algorithm works like a nested fixed point problem. The first part of Proposition (1.1) solves for the aggregate equilibrium, while the second part represents a nested subroutine which solves for the modal transport flows conditional on a guess for the market access terms (Proposition (1.2)). A computational benefit arises because - given a candidate solution for the aggregate equilibrium - the transport equilibrium is entirely disentangled from the determination of the aggregate spatial equilibrium, while the aggregate equilibrium inherits the computational properties in terms of speed and scalability of a (gravity) spatial equilibrium system.

From a practitioner's point of view, the proposition implies that given observed traffic flows on the primary network, bilateral flows on the secondary network[28], $\left(\Xi_{ik}, \Xi_{ik,m}\right)$ as well as knowledge of the model parameters, $\{\alpha, \beta, \theta, \lambda_1, \lambda_m, \eta\}$, we can employ the model to evaluate infrastructure improvements along the primary network or at terminal stations, thus improving the connectedness of the primary and secondary transport network. We further extend the model to allow for a partitioned set of domestic and foreign locations with international trade between a subset of domestic nodes and foreign locations, as presented in Appendix A.10.

Relative to the existing literature, the proposition provides a straightforward extension of equation (28) and (29) in AA2022. The only difference is the presence of a nested subroutine, which given a candidate solution for spatial equilibrium, $\{P_i, \Pi_i\}$, solves the transport equilibrium at the edge-level, i.e. the set of equations defined by Equation (22). This implicitly determines the modal choice taking into account endogenous congestion and generates an endogenous aggregate transport cost.[29] This adjustments adds a novel channel towards evaluation infrastructure improvements. In this setting, the impact of improving transport infrastructure has the same direct and general equilibrium effect as in AA2022 where route choice is impacted, congestion can be alleviated, input and output markets can adjust and where all this adds up to welfare gains. In our setting, additionally, we also feature a direct interplay between the primary and secondary network. Mode-specific infrastructure investments can lead to modal diversion and thus alleviate congestion on the alternative transport network. The extent to which this might occur depends on the cross-sectional variation in the access to the secondary transport system, which is reflected by variations in the weights on the final term across space.

5.2 Model Calibration to US Multimodal Network

In order to calibrate our framework to the US multimodal network, we require data on income, population, as well as geo-spatial and traffic data on roads, rail, waterways, ports, and inland intermodal terminals. We utilize income, population, as well as road geo-spatial and traffic data from AA2022. The additional rail, waterways, ports, and intermodal terminals data have been introduced in Section 2.1. To allow

[28] Notice the slight abuse of notation here. While $\Xi_{ij,1}$ refers to the edge-specific traffic along the primary network, $\Xi_{ij,m}$ instead refers to modal flows between i and j along the secondary networks and is therefore not necessarily edge-specific. However, $\Xi_{ij,m}$ summarizes modal traffic in the sense that it refers to any flows between i and j no matter their origin or destination on the primary network. This is convenient - as we will argue below - since this is the data moment that is directly observed in the rail traffic data.

[29] We can *explicitly* characterize the change in the equilibrium traffic flows along the primary and secondary transport system. Given the equilibrium changes in market access terms $\left(\hat{P}_i, \hat{\Pi}_i,\right)$, as well as the equilibrium changes of mode and aggregate transport costs,$\left(\hat{t}_{ik}, \hat{t}_{ik,m},\right)$, the resulting change in the traffic flows can be computed using the following formulae:

$$\begin{aligned} \hat{\Xi}_{ik} &= \hat{t}_{ik}^{-\theta} \times \hat{P}_i^{-\theta} \times \hat{\Pi}_k^{-\theta} \\ \hat{\Xi}_{ik,m} &= t_{ik,m}^{-\eta} \times \hat{t}_{ik}^{\theta-\eta} \times \hat{P}_i^{-\theta} \times \hat{\Pi}_k^{-\theta} \end{aligned} \tag{23}$$

which allows us to account for the changes in the observed traffic flows. This can be a convenient tool to analyse the implied traffic and environmental impact of infrastructure investments.

for international trade we obtain Census Bureau data on flows from international locations to domestic ports.[30] Our resultant sparse graph representation of the multimodal transport network collapses the high-dimensional geo-spatial information from each mode-specific shapefiles and retains nodes at endpoints or intersections. Cities containing at least 10,000 people (CBSAs) are also represented by a single node within the network. The resulting graph comprises of 228 nodes and and 704 edges.

In a last step before the counterfactual analysis, we discuss our choice of the model parameters $\{\alpha, \beta, \theta, \lambda_1, \lambda_m, \eta\}$. For the local productivity spillover α, local amenity spillovers β, and shape parameter θ, we follow the literature (Allen and Arkolakis, 2022) and set $\theta = 8$, $\alpha = .1$ and $\beta = -.3$. For the road congestion elasticity, we use $\lambda_1 = 0.092$ following AA2022 who estimates this parameter by regressing the observed speed on individual highway segments against a measure of instrumented traffic.

The last two parameters come from our own estimates. For the intermodal terminal congestion parameter λ_m, we first determine the impact of port traffic on the observed duration a ship spends at port using an instrument which measures the trade exposure at these ports (Table 2, Section 4.2). We use $\lambda_m = 0.096$ after converting this time elasticity into a cost measure, relying on the ad valorem tariff estimate on an additional day in transit from Hummels and Schaur (2013). For the modal elasticity of substitution η, we estimate the impact of road infrastructure improvements on rail relative to road traffic flows using the instruments from Duranton and Turner (2011, 2012); Duranton, Morrow and Turner (2014). We use $\eta = 1.099$ (Table 1, Section 4.1). The non-overlapping confidence intervals for η and θ highlight a significant difference between them, supporting the validity of our modeling approach which allows for two distinct elasticities to govern transport mode and route-sourcing decisions separately. As discussed earlier in Section 3.4, this also suggests that a simple extension of the AA2022 to allow for mode-specific subnetworks is insufficient to accurately capture both decisions (footnote 8).

5.3 Predicted versus Observed Transport Mode-level Trade Flows

In order to assess the ability of our model to capture actual observed trade flows for different transport modes, we compare our model predictions for mode-specific trade flows between CBSAs against commensurate data from the Commodity Flow Survey (CFS). The model predicted mode-specific trade flows are calculated using Equation (26) and observed mode-specific traffic flow data from the CFS.[31] We include all three transport modes—truck, rail, and barges. Figure 6a shows the scatter plot between the observed and predicted mode-specific trade flows, conditional on origin and destination fixed effects. This allows for a pair-wise comparison between the observed and predicted mode-specific flows at the origin-destination-level, independent of origin or destination characteristics. We find a strong positive correlation of 0.75 with all three modes, indicating that the model predictions reflect observed mode-

[30] To limit the dimensionality of the simulation, we group the flows into five geographic regions: Africa, Asia, Australia and Oceania, Europe, North America and South/Central America.

[31] See Online Appendix A.9 for detailed derivations of the model predicted mode-specific trade flows.

specific trade flows. These results are also an improvement in terms of model fit relative to previous literature focusing on one mode. As a comparison point, AA2022 finds a strong positive correlation of 0.60 using truck flows. Column (1) of Table A.10 reports the univariate regression outcome between the residualized observed and predicted trade flows, weighted by trade weight in tons. We find a significantly positive relationship, with a coefficient of 1 within one standard error.

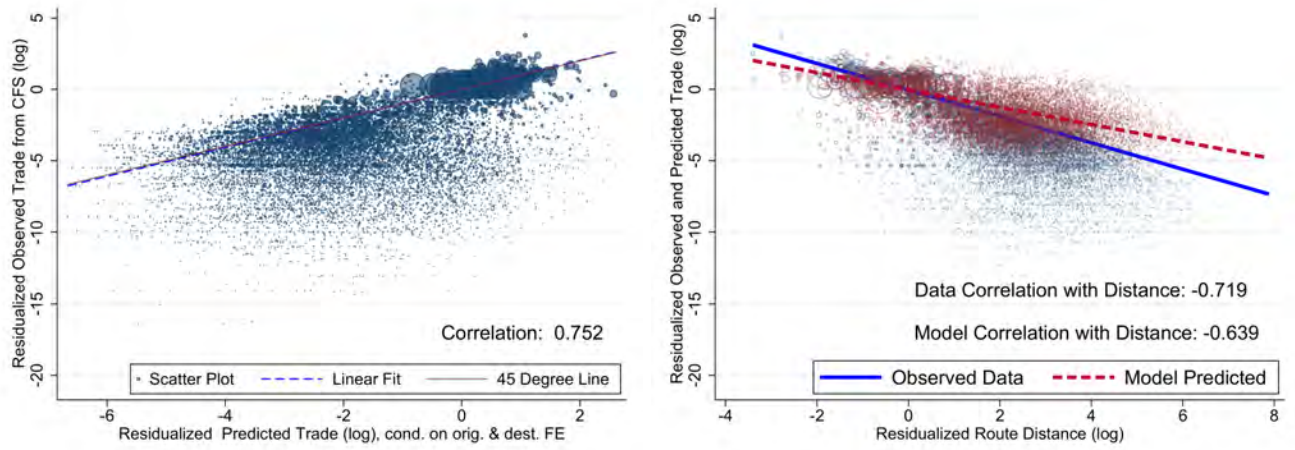
Additionally, we can evaluate the validity of our model by examining its gravity model implications. The standard gravity model implies a negative log-linear relationship between trade and trade costs at the bilateral level, both internationally between countries (Head and Mayer, 2014) and domestically between cities within a country (Allen and Arkolakis, 2014; Fajgelbaum and Schaal, 2020). Following previous literature, we compare the relationship between observed trade flows and distance to the relationship between the predicted trade flows and distance, similarly conditioned on origin and destination fixed effects as above. Figure 6b find similar strong negative correlations between the two, strong negative correlations of 0.64 in the model and 0.72 in the data. We find significantly negative relationships in linear regressions for both, with a negative coefficient of 0.61 in the model and 0.93 in the data. The similarity between the predicted and observed relationships with distance strengthens the validity of our model, suggesting that it provides reasonable approximations for the distribution of trade flows.

Looking into specific transport modes, we find strong positive correlations of 0.81 and 0.56 between the residualized observed and predicted trade flows for trucks and multiple modes respectively. Column (2) of Table A.10 reports the regression outcome between the residualized observed and predicted trade flows for the three transport modes, weighted by trade weight in tons. We find a significantly positive relationship that is close to 1 for all three modes, with the rail coefficient slightly higher, suggesting that observed rail flows exceed model predictions. This may be due to the CFS data capturing flows for all commodities, while our model only predicts containerized flows. From our confidential waybill data, containerized trade accounts for 43 percent of all rail car loads in 2012.

5.4 Unimodal versus Multimodal Network: Highway System Improvements

As mentioned in Section 3.2, our model is isomorphic to the unimodal network in AA2022 when the mode choice elasticity is assumed to be equal to the route choice elasticity. This allows us to estimate the same highway network improvement to compare our multimodal network results to a unimodal one. We find a high correlation of 0.76 and relative rankings that are consistent with their results (Figure A.5). However, our welfare estimates are higher because of the positive spillovers onto other transport modes in our multimodal network. Road improvements have the propensity to lower both the effects of congestion on roads as well as at intermodal terminals.

Figure 6. Model Fit: Predicted and Observed Trade Flows



(a) Predicted and Observed Bilateral Trade Flows **(b)** Trade Flows versus Distance

Notes: The figure in Panel (a) visualizes the comparison between observed bilateral origin to destination mode-specific trade flows with the mode-specific trade flows predicted by the multimodal economic geography model based on observed traffic data along the transport network. The observed mode-specific trade flows between metropolitan areas on the y-axis are obtained from the 2012 Commodity Flow Survey (CFS) data while the predicted mode-specific trade flows are on the x-axis. Both the observed and predicted trade flow measures are in logs and residualized using origin and destination fixed effects. This allows for the comparison to come from similarities at the origin-destination pair-level. The figure in Panel (b) visualizes the observed and predicted bilateral trade flows against distance, all residualized using origin and destination fixed effects. In both panels, circle sizes indicate trade weight in tons.

6 Multimodal Network Improvements and Disruptions

In this section, we apply our framework to evaluate the economic and environmental impacts of targeted infrastructure investments and policy-relevant scenarios, taking the multimodal nature of the US domestic transport system into account.[32] Our main application compares the welfare impact of investing in different terminals across the US, improving the intermodal integration of the multimodal transport network. Additionally, we quantify the welfare changes and environmental impacts of three policy-relevant scenarios: (1) the impact of losing rail network access, motivated by the averted rail strike in the fall of 2022, (2) the impact of repealing the Merchant Marine Act of 1920 (also known as the Jones Act), which restricts ships transporting goods between US ports to those owned, crewed, and built by US citizens, and (3) the impact of disruptions to the Panama Canal due to drought conditions.

6.1 Gains from Intermodal Terminal Improvements

While previous papers have focused on estimating the welfare effects of improving individual segments of the US road network and particularly highways, less is known about the welfare impact of improving the degree to which the US multimodal transport network is interconnected. In order to evaluate this, we will use the counterfactual equations of Proposition (1) to estimate the aggregate welfare impact

[32] We explicitly embrace a computational approach in this section, employing the 'hat algebra' developed in Section 5.1. Alternatively, one could potentially derive sufficient statistics that characterize welfare elasticities in terms of traffic flows as in Office of the State Auditor (2019) and in principle the recursive formulation of the equilibrium condition (Equation (4)) might be a convenient aid in doing so as highlighted in our discussion in Subsection 3.6. However, the presence of rich a set of congestion elasticities on the primary mode and intermodal terminals makes this approach less desirable.

($\hat{W} = \hat{\chi}^{-\frac{1}{\theta}}$) of a small (1%) improvement to the switching cost, $s_{ii,m}$, at each intermodal node across the system, given the calibrated parameters and observed traffic on road, rail, waterways, and ports.

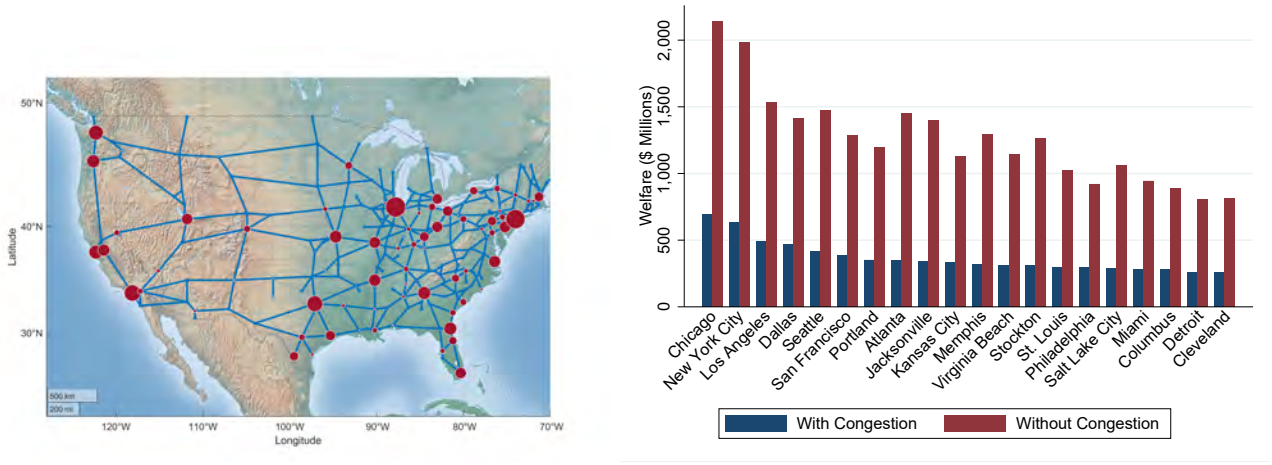
Our results are visualized in Figure 7a and the top 10 terminals with the highest impact are listed in Table 3. For the welfare benefits, we convert the welfare elasticity into a dollar amount using a compensating variation approach by calculating the necessary increase, in millions of chained 2012 US dollars, in annual US real GDP (valued at $19 trillion) to achieve the same welfare improvement that we estimate. Unsurprisingly, some of the intermodal terminals that generate the largest welfare impacts are populous coastal cities like New York City and Los Angeles (which also includes Long Beach), which both rank top two and three respectively (Table 3). However, many of the most impactful terminals are located in center of the United States like Chicago, Dallas, and Atlanta—highlighting the role of multimodal network transporting goods from coastal regions to the interior. Chicago, the top ranked terminal, has the highest rail throughput indicating its central role as a intermodal hub for the Midwest (Table 3, Column (4)). Several of these terminals, like Dallas, Kansas City, and Atlanta, do not have large throughput flows or a lot of terminals (Table 3, Column (3)). However, their importance lies in their strategic geographical position along the multimodal network that facilitates the connection between coastal areas and the interior regions (Figure 7a).

In contrast, link improvements within the unimodal highway network have identified gains in short coastal segments linking densely populated areas (like Boston-Philadelphia and Los Angeles-San Diego), and in trade thoroughfares via Indiana (AA2022). Our terminal improvement gains instead are mostly in the central US, highlighting the role of multimodal transportation taking place over longer distances and linking coastal to interior regions (as indicated in Figure 1). These locations are highly central to the transportation system and represent potential bottlenecks. The implied welfare benefit of lowering the switching costs in the most central nodes could represent a welfare gain equivalent to increasing US GDP between 300-700 million USD (in 2012 USD, Table 3, Column (6)).

We further show that congestion plays an important role in determining the magnitudes of these welfare effects. Specifically, the benefits from improvements without congestion are 2.1 times higher at Chicago relative to improvements with congestion, 2 times for Dallas, 2.4 times at Kansas City, and 3.2 times for Atlanta (Figure 7b). The welfare benefits are on average 2.5 times higher without congestion for the top 50 terminals compared to benefits with congestion (Figure A.6). Targeting infrastructure investments at nodes that are likely bottlenecks is crucial since these bottlenecks tend to occur due to limited capacity or a lack of alternative routes, both of which can hinder the transportation of goods. Both these factors are particularly pertinent for intermodal terminals, suggesting that alleviating congestion at terminals can yield large positive returns for the overall performance of the multimodal network.

Additionally, we combine our welfare benefits with cost assessments in order to compute the return on investment (ROI) for each of these terminal investments. To operationalize this, we utilize equation (42) in AA2022 adapted for terminal operations to determine the amount of container volume required

Figure 7. Welfare Benefits of Improving Intermodal Terminal



(a) Map of Intermodal Terminals **(b)** Welfare Benefits with and without Congestion

Notes: Panel (a) visualizes the welfare impact of lowering the transshipment cost in each node by 1 percent. The larger the dots, the larger the welfare gains. The blue lines indicate the graph representation of the primary road network. State boundaries are included. See Table 3 for details on the top 10 list of intermodal terminals with the highest welfare impacts. Panel (b) presents a comparison of the welfare impacts for the top 20 most impactful terminals with and without congestion effects. See Figure A.6 for a scatterplot comparing both effects for all terminals.

Table. 3. Ranking: Welfare Benefits of Improving Intermodal Terminals

	(1) CBSA Name	(2) Population	(3) Terminals	(4) Throughput	(5) ROI	(6) Benefit ($m)	(7) Cost ($m)
1	Chicago	9368268	88	3456228	0.249	691	553
2	New York City	14745610	29	497852	1.820	635	225
3	Los Angeles	9639715	38	2278880	-0.624	490	1301
4	Dallas	4513776	13	564160	4.581	465	83
5	Seattle	2189215	20	644052	1.917	418	143
6	San Francisco	3863536	14	104312	2.757	384	102
7	Portland	1641801	30	141432	8.891	352	36
8	Atlanta	1627623	28	610280	3.237	347	82
9	Jacksonville	936317	18	265960	6.995	341	43
10	Kansas City	1767872	55	362920	7.225	333	40

Notes: The table shows the ten terminals where a one percent reduction of the transshipment cost generates the highest benefit. Column (1) indicates the core based statistical areas (CBSA) name of the node, which includes both metropolitan and micropolitan areas. The terminal's population, number of terminals, and rail throughput in TEUs are reported in Columns (2), (3), and (4) respectively. Column (5) shows the imputed return on investment (ROI), Column (6) calculates how much 2012 US GDP would need to increase in order to match the overall welfare gain, while Column (6) presents the required cost of making this one percent cost decrease. For an extended version including the top 30 terminals see Table A.11.

to achieve a one percent reduction in switching cost at these terminals.[33] Next, we calculate the number of gantry cranes in container terminals required to achieve this higher container volume using industry estimates. Since these terminals require land, each location will have a different construction cost based on local property prices which we obtain from Albouy, Ehrlich and Shin (2018). Assuming a 20-year depreciation schedule, a 5% annual maintenance cost, and a 3% borrowing cost per existing literature, we estimate that 10% of the construction cost is accrued on an annual basis. We estimate high terminal improvement costs for populous and coastal cities like Los Angeles and New York City (Table 3, Column (7)). These cost estimates translate into varying levels of ROIs for these terminals (Table 3, Column

[33]This implies a (log) proportional relationship between number of terminals and switching costs, i.e. $\hat{s}_{kk,m} = \hat{terminals}_{kk,m}^{-\lambda_m}$. where $\hat{s}_{kk,m}$ denotes the mode specific (log) change in the switching cost and $\hat{terminals}_{kk,m}$ denotes the necessary (log) change in terminal capacity.

(5)). We find low ROIs for large cities like Chicago and New York City, and even negative returns for Los Angeles. On the other hand, we estimate high ROIs smaller locations that are geographically central like Kansas City, Jacksonville, Dallas, and Portland.

In order to calculate the environmental impact from this counterfactual, we obtain data on transport mode-specific greenhouse gas (GHG) emissions from EPA (2022*a*). For each counterfactual improvement of the intermodal terminals, we calculate the change in mode-specific traffic flows using Equation (23) and multiply this change with the mode-specific GHG emissions in order to determine the change in GHG emissions.[34] We employ EPA (2022*b*)'s estimates when calculating the social cost of GHG emissions.

Table 4 reports the environmental impacts of the intermodal terminal improvements. For all terminals, this decrease in the cost of accessing multimodal transport raises the relative cost of using truck transport which decreases road traffic flows, resulting in a decrease in GHG emissions from trucks—highlighting the modal substitution effect (Table 4, Column (4)). Since the switching cost of intermodal terminals have decreased, rail traffic flows will increase in response which increases GHG emissions for rail (Table 4, Column (5)). However, since rail is much less polluting than road, the overall GHG emission change is net decrease for all the terminals. The implied welfare benefit of lowering the switching costs in some of the most central nodes could result in an unintended environmental benefit, equivalent to between 23 to 45 million USD (Table 4, Column (6)). Our findings are align with the objectives of the inter-agency US National Blueprint for Transportation Decarbonatization, which prioritizes improving multimodal freight transport, and the European Green Deal, which aims to shift 75% of inland freight from road to rail and inland waterways (DOE, DOT, EPA, and HUD, 2024; European Commission, 2021).

6.2 Impact from Losing Rail Network Access

The potential of a rail strike loomed in the fall of 2022, only averted due to intervention from US President Biden (Kanno-Youngs and Cochrane, 2022). A strike by rail workers which would have shut down the entire rail transport system. In order to evaluate the impact of losing access to the rail network, we estimate the aggregate welfare impact of increasing the cost on all rail links to prohibitively high levels (twenty-fold). The first row in Table 5 shows that the welfare cost of losing the rail network is equivalent to decreasing US GDP by about 230 billion USD (in 2012 USD). While existing literature have quantified the importance of the road and highway transport network, the value of the rail network has received less attention, except in historical analyses. Our estimate of the value of the rail network is roughly half the value of the US highway system as estimated by Jaworski, Kitchens and Nigai (2023), and about 40% more than the historic adjusted value of railroads to the agriculture sector according to Donaldson and Hornbeck (2016).

[34]The aggregate growth from these improvements outweighs the partial equilibrium effect from mode-specific traffic changes and leads to a global increase in greenhouse gas emissions. In order to isolate the partial equilibrium effect, we keep overall (GHG weighted) traffic constant.

Table. 4. Ranking: Environmental Impact of Intermodal Terminal Improvements

	(1) CBSA Name	(2) Benefit ($m)	(3) Cost ($m)	(4) Truck GHG (kt)	(5) Rail GHG (kt)	(6) GHG Benefit ($m)
1	Chicago	691	553	-148.62	19.15	45.33
2	New York City	635	225	-148.09	19.05	45.13
3	Los Angeles	490	1301	-107.17	13.36	32.64
4	Dallas	465	83	-89.36	11.82	27.23
5	Seattle	418	143	-101.76	12.92	31.00
6	San Francisco	384	102	-100.82	12.97	30.73
7	Portland	352	36	-72.86	9.54	22.21
8	Atlanta	347	82	-82.83	11.12	25.24
9	Jacksonville	341	43	-81.10	10.37	24.74
10	Kansas City	333	40	-76.64	9.89	23.37

Notes: The table shows the environmental impact from the ten terminals where a one percent reduction of the transshipment cost generates the highest benefit. Column (1) indicates the core based statistical areas (CBSA) name of the node, which includes both metropolitan and micropolitan areas. Column (2) calculates how much 2012 US GDP would need to increase in order to match the overall welfare gain (reproduced from Table 3 Column (6)), while Column (3) presents the required cost of making this one percent cost decrease (reproduced from Table 3 Column (7)). Column (4) shows the change in truck GHG emissions which comes from the change in road traffic flows. Column (5) shows the change in rail GHG emissions which comes from the change in rail traffic flows. Column (6) presents the net social benefit from the changes in mode-specific GHG emissions. Waterway emissions are omitted for brevity.

Table. 5. Welfare Impact of Policy Scenarios: The Role of Congestion

	(1) Scenario	(2) Benefit ($bn)	(3) Benefit without Terminal Congestion ($bn)	(4) Benefit without All Congestion ($bn)
1	Railroad Strike	-230.46	-236.40	-278.94
2	Removal of the Jones Act	3.15	11.73	16.16
3	Panama Canal	-2.67	-7.64	-10.29

Notes: The table shows the welfare impact each of the three scenarios indicated by each row. Column (1) labels each scenario. Column (2) calculates how much 2012 US GDP would need to change in order to match the overall welfare change from each scenario. Column (3) is calculated the same way as Column (2) except with the removal of congestion on intermodal terminals ($\lambda_m = 0$). Column (4) is calculated the same way as Column (2) except with the removal of congestion on intermodal terminals ($\lambda_2 = 0$) and roads ($\lambda_1 = 0$).

How much does congestion affect these results? We next estimate the same increase in rail link costs, but reducing the congestion at intermodal terminals to zero ($\lambda_m = 0$). We find a very small increase in the negative welfare impact of 2.5%, which is expected since the terminal congestion effects are not binding due to prohibitively high cost of rail—there is very little traffic flows on the rail network. The small increase is likely due to intermodal terminals that are also ports. When we remove all congestion effects, however, we see a 21% larger negative impact which is equivalent to decreasing US GDP by about 279 billion USD (in 2012 USD, first row of Table 5, Column (4)). As seen from our terminal improvement counterfactuals, congestion has a compounding effect on welfare (Figures 7b and A.6).

Next, we evaluate the environmental impact of the rail strike. The rail strike scenarios pushes the cargo that would have gone via rail onto road instead, increasing the truck GHG emissions by a significant amount (first row of Table 6, Column (2)). Even though rail GHG emission decrease due to the lack of rail transport (first row of Table 6, Column (3)), trucks are much more polluting. As a result, the modal substitution towards trucks due to the rail network loss results in a huge decrease in the social cost of GHG by 12 billion dollars (first row of Table 6, Column (4)). Overall, losing access to the rail network would not just lead to a substantial decline in welfare but also cause significant environmental harm due to the modal substitution towards truck transportation.

Table. 6. Environmental Impact of Policy Scenarios

	(1) Scenario	(2) Truck GHG Change (kt)	(3) Rail GHG Change (kt)	(4) GHG Benefit ($bn)	(5) Benefit ($bn)
1	Railroad Strike	38947	-5171	-11.88	-230.46
2	Removal of the Jones Act	-589	-47	0.19	3.15
3	Panama Canal	1524	111	-0.45	-2.67

Notes: The table shows the environmental impact of each of the three scenarios indicated by each row. Column (1) labels each scenario. Column (2) shows the change in truck GHG emissions which comes from the change in road traffic flows. Column (3) shows the change in rail GHG emissions which comes from the change in rail traffic flows. Column (4) presents the net social cost or benefit from the changes in mode-specific GHG emissions. Column (5) calculates how much 2012 US GDP would need to change in order to match the overall welfare change from each scenario (reproduced from Table 5 Column (2)). Waterway emissions are omitted here for brevity.

6.3 Repeal of Jones Act

Next, we examine the consequences of repealing the Jones Act, the Merchant Marine Act of 1920. This is an active US trade regulation which requires all ships transporting freight and passengers between US ports to be built in the US, owned by US citizens, and mostly crewed by US citizen. Jones Act ships on average cost 2.7 times to operate than foreign-flag equivalents (MARAD, 2011). We evaluate the repeal of the Jones Act by decreasing the link cost of waterways by this MARAD (2011) estimate. The second row in Table 5 shows that the welfare gain of repealing the Jones Act and allowing foreign ships to transport cargo between US ports is equivalent to increasing US GDP by about 3.2 billion USD (in 2012 USD). Our estimate is towards the lower end of the range of estimates by the USITC: $2.8 billion (USITC, 1995) to $9.8 billion (USITC, 1991). There are a few reasons that contribute to our estimate being lower. First, we estimate this impact only for the continental US. The Jones Act repeal will have much larger benefits for Hawaii, Alaska, and Puerto Rico since these locations rely a lot more on maritime transport and have less modal substitution options. Secondly, our estimates are based on the the US multimodal network today. The Jones Act, established in 1920, is more than 100 years old. Over this long time period, the multimodal network has adjusted to substitute for the expensive waterway transport. As an example, Figure 3 shows that there is rail and road networks up and down the coast lines. Had waterways been less expensive, these networks might have been potentially built further inland.

In this scenario, the congestion effects of intermodal terminals play a much larger role in informing the welfare estimates. Without intermodal terminal congestion, the welfare benefits from repealing the Jones Act increase by 272% (second row in Table 5 Column (3)). This is due to the gains from being able to increase waterway transport use once it is cheaper. Without all congestion (road and terminals), the welfare gains from the repeal increase by a further 38 percent (second row in Table 5 Column (4)).

The repeal of the Jones Act decreases emissions overall. This is because cargo would substitute towards using waterways due to the lower cost, and so road and rail traffic would decrease as a result (second row of Table 6, Columns (2) and (3)). Since barge emissions are much lower, this results in a GHG benefit increase of 0.2 billion dollars (second row of Table 6, Column (4)).

6.4 Panama Canal Disruptions

Our third scenario studies the impact of the drought conditions in the Panama Canal, which decreases the number of ships that can access it (Rojanasakul, 2024). To trace out the importance of the Panama Canal and how it interacts with the US multimodal transportation system, we include in our calibration a (coarse) notion of routed imports at the port-level. Specifically, we include linkages from ports to foreign destinations and calibrate the importance of these linkages with port-destination specific trade flows.[35] To simulate the impact of the Panama Canal, we assume that routes that typically require usage of the Panama Canal (e.g. Asia to ports on the East Coast) are affected by transport cost increase. We evaluate this disruption by increasing the cost of accessing the canal by five times.

The third row in Table 5 shows that the welfare cost of disruptions at the Panama Canal is equivalent to decreasing US GDP by about 2.7 billion USD (in 2012 USD). To the best of our knowledge, this is the first estimate of this disruption allowing for modal and route substitution, including at ports. Here, removing congestion again compounds these effects. Without terminal congestion, the effects increase by more than 186 percent. This large increase is driven by the modal substitution from water to rail and road, and the intermodal terminals playing a crucial role in facilitating this transfer. Without terminal and road congestion, the increase is a further 35 percent.

The modal substitution due to the canal disruptions in turn increases the truck and rail traffic flows (which increases GHG emissions on both modes (third row of Table 6). This results in an overall decrease in the GHG benefit of 0.5 billion dollars (third row of Table 6, Column (4)).

7 Conclusion

We examine the economic and environmental impacts of improvements and disruptions in multimodal transport networks. We focus on how these outcomes will depend on the geography of the multimodal transport network, the placement of intermodal terminals that allow for switches between modes of transport, as well as the relative cost of transport across modes. By incorporating these features, we provide a framework to realistically evaluate infrastructure policies and disruptions across the entire multimodal transport network. Our paper highlights that strategic investments in intermodal terminals—which can be potential bottlenecks—can enhance the robustness and efficiency of transport networks, supporting global and local trade as well as economic stability (Ganapati and Wong, 2023). Modal substitution from road to rail traffic result in added environmental benefits, supporting the objectives of the US National Blueprint for Transportation Decarbonization, which focuses on improving multimodal freight transport, and the European Green Deal, which seeks to shift 75% of EU inland freight from road to rail and inland waterways (DOE, DOT, EPA, and HUD, 2024; European Commission, 2021).

[35] We collect data of destination specific imports from the Census Bureau and aggregate across five major regions: Central America, Africa, Asia, Australia and Oceania, Europe, North America (excl USA), South/Central America.

References

Akbar, Prottoy A, Victor Couture, Gilles Duranton, and Adam Storeygard. 2023. "The fast, the slow, and the congested: Urban transportation in rich and poor countries." National Bureau of Economic Research.

Albouy, David, Gabriel Ehrlich, and Minchul Shin. 2018. "Metropolitan land values." *Review of Economics and Statistics*, 100(3): 454–466.

Allen, Treb, and Costas Arkolakis. 2014. "Trade and the Topography of the Spatial Economy." *The Quarterly Journal of Economics*, 129(3): 1085–1140.

Allen, Treb, and Costas Arkolakis. 2022. "The Welfare Effects of Transportation Infrastructure Improvements." *The Review of Economic Studies*.

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

Antràs, Pol, and Alonso Gortari. 2020. "On the geography of global value chains." *Econometrica*, 88(4): 1553–1598.

Baum-Snow, Nathaniel. 2007. "Did highways cause suburbanization?" *The quarterly journal of economics*, 122(2): 775–805.

Beuthe, Michel, Bart Jourquin, and Natalie Urbain. 2014. "Estimating freight transport price elasticity in multi-mode studies: A review and additional results from a multimodal network model." *Transport Reviews*, 34(5): 626–644.

Bonadio, Barthélémy. 2021. "Ports vs. Roads: Infrastructure, Market Access and Regional Outcomes." Working Paper.

Boyles, Stephen D., Nicholas E. Lownes, and Avinash Unnikrishnan. 2021. *Transportation Network Analysis.* Vol. 1. 0.89 ed.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou. 2020. "Geography, transportation, and endogenous trade costs." *Econometrica*, 88(2): 657–691.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou. 2024. "Investment in Infrastructure and Trade: The Case of Ports." National Bureau of Economic Research.

Brooks, Leah, Nicolas Gendron-Carrier, and Gisela Rua. 2018. "The local impact of containerization." *Finance and Economics Discussion Series*, 45.

Carballo, Jeronimo, Alejandro Graziano, Georg Schaur, and Christian Volpe Martincus. 2021. "Import Processing and Trade Costs." CESifo Working Paper no. 9170.

CBO. 2022. "Emissions of Carbon Dioxide in the Transportation Sector." Congressional Budget Office Publication 58566.

Census Bureau. 2003-2021. "USA Trade Online." *https://usatrade.census.gov*, Last accessed on August 2024.

Coşar, A Kerem, and Banu Demir. 2018. "Shipping inside the box: Containerization and trade." *Journal of International Economics*, 114: 331–345.

Couture, Victor, Gilles Duranton, and Matthew A Turner. 2018. "Speed." *Review of Economics and Statistics*, 100(4): 725–739.

Cristea, Anca, David Hummels, Laura Puzzello, and Misak Avetisyan. 2013. "Trade and the greenhouse gas emissions from international freight transport." *Journal of environmental economics and management*, 65(1): 153–173.

Cuturi, Marco. 2013. "Sinkhorn Distances: Lightspeed Computation of Optimal Transportation Distances." *arXiv [stat.ML]*.

Daly, Andrew, and Michel Bierlaire. 2006. "A general and operational representation of Generalised Extreme Value models." *Trans. Res. Part B: Methodol.*, 40(4): 285–305.

Degiovanni, Pedro, and Ron Yang Yang. 2023. "Economies of Scale and Scope in Railroading." Working Paper.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum. 2008. "Global Rebalancing with Gravity: Measuring the Burden of Adjustment." National Bureau of Economic Research, Inc NBER Working Papers 13846.

Djankov, Simeon, Caroline Freund, and Cong S Pham. 2010. "Trading on time." *The review of Economics and Statistics*, 92(1): 166–173.

DOE, DOT, EPA, and HUD. 2024. "US National Blueprint for Transportation Decarbonatization." *https://www.energy.gov/sites/default/files/2023-01/the-us-national-blueprint-for-transportation-decarbonization.pdf*, US Department of Energy, US Department of Transportation, US Environmental Protection Agency, US Department of Housing and Urban Development.

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

Ducruet, César, Réka Juhász, Dávid Krisztián Nagy, and Claudia Steinwender. 2020. "All aboard: The effects of port development." National Bureau of Economic Research.

Dunn, Jason, and Fernando Leibovici. 2023. "Navigating the Waves of Global Shipping: Drivers and Aggregate Implications." Federal Reserve Bank of St. Louis, Research Division.

Duranton, Gilles, and Diego Puga. 2023. "Urban growth and its aggregate implications." *Econometrica*, 91(6): 2219–2259.

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

Duranton, Gilles, and Matthew A Turner. 2012. "Urban growth and transportation." *Review of Economic Studies*, 79(4): 1407–1440.

Duranton, Gilles, Peter M Morrow, and Matthew A Turner. 2014. "Roads and Trade: Evidence from the US." *Review of Economic Studies*, 81(2): 681–724.

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

EPA. 2022*a*. "Inventory of US Greenhouse Gas Emissions and Sinks: 1990-2020." *https://www.epa.gov/ghgemissions/draft-inventory-us-greenhouse-gas-emissions-and-sinks-1990-2020*, US Environmental Protection Agency, EPA 430-R-22-0003.

EPA. 2022*b*. "Report on the social cost of greenhouse gases: Estimates incorporating recent scientific advances." US Environmental Protection Agency Washington, DC.

EPA. 2024. "Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2022." U.S. Environmental Protection Agency, EPA 430-R-24-004.

European Commission. 2021. "Sustainable and Smart Mobility Strategy." *https://transport.ec.europa.eu/system/files/2021-04/2021-mobility-strategy-and-action-plan.pdf*, Directorate-General for Mobility and Transport.

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

Fajgelbaum, Pablo D, Cecile Gaubert, Nicole Gorton, Eduardo Morales, and Edouard Schaal. 2023. "Political Preferences and the Spatial Distribution of Infrastructure: Evidence from California's High-Speed Rail." National Bureau of Economic Research.

Fan, Jingting, and Wenlan Luo. 2020. "A Tractable Model of Transshipment." Working Paper.

Fan, Jingting, Yi Lu, and Wenlan Luo. 2019. "Valuing domestic transport infrastructure: a view from the route choice of exporters." *The Review of Economics and Statistics*, 1–46.

Galichon, Alfred. 2021. "The unreasonable effectiveness of optimal transport in economics." *arXiv [econ.GN]*.

Ganapati, Sharat, and Woan Foong Wong. 2023. "How far goods travel: global transport and supply chains from 1965–2020." *Journal of Economic Perspectives*, 37(3): 3–30.

Ganapati, Sharat, Woan Foong Wong, and Oren Ziv. 2021. "Entrepot: Hubs, scale, and trade costs." National Bureau of Economic Research Working Paper 29015.

Gibbons, Stephen, Stephan Heblich, and Edward W Pinchbeck. 2024. "The spatial impacts of a massive rail disinvestment program: The Beeching Axe." *J. Urban Econ.*, 143(103691): 103691.

Harrigan, James. 2010. "Airplanes and comparative advantage." *Journal of International Economics*, 82(2): 181–194.

Head, Keith, and Thierry Mayer. 2014. "Gravity equations: Workhorse, toolkit, and cookbook." In *Handbook of international economics*. Vol. 4, 131–195. Elsevier.

Heiland, Inga, Andreas Moxnes, Karen-Helene Ulltveit-Moe, and Yuan Zi. 2019. "Trade From Space: Shipping

Networks and The Global Implications of Local Shocks." C.E.P.R. Discussion Papers CEPR Discussion Papers 14193.

Horn, Roger A, and Charles R Johnson. 2012. *Matrix analysis*. Cambridge university press.

Hummels, David. 2007. "Transportation Costs and International Trade in the Second Era of Globalization." *J. Econ. Perspect.*, 21(3): 131–154.

Hummels, David L, and Georg Schaur. 2013. "Time as a trade barrier." *American Economic Review*, 103(7): 2935–2959.

Jaworski, Taylor, Carl Kitchens, and Sergey Nigai. 2023. "Highways and globalization." *International Economic Review*, 64(4): 1615–1648.

Kanno-Youngs, Zolan, and Emily Cochrane. 2022. "Biden Signs Legislation to Avert Nationwide Rail Strike." *https://www.nytimes.com/2022/12/02/us/politics/rail-strike-biden.html*, New York Times.

Kitthamkesorn, Songyot, Anthony Chen, and Xiangdong Xu. 2015. "Elastic demand with weibit stochastic user equilibrium flows and application in a motorised and non-motorised network." *Transportmetrica A: Transport Science*, 11(2): 158–185.

Lind, Nelson, and Natalia Ramondo. 2023. "Trade with Correlation." *American Econonomic Review*, 113(2): 317–353.

Li, Xinyan, Chi Xie, and Zhaoyao Bao. 2022. "A multimodal multicommodity network equilibrium model with service capacity and bottleneck congestion for China-Europe containerized freight flows." *Transportation Research Part E: Logistics and Transportation Review*, 164: 102786.

Lugovskyy, Volodymyr, Alexandre Skiba, and David Terner. 2022. "Unintended Consequences of Environmental Regulation of Maritime Shipping: Carbon Leakage to Air Shipping." Working Paper.

MARAD. 2011. "Comparison of US and Foreign-flag Operating Costs." US Department of Transportation, Maritime Administration.

McFadden, Daniel. 1977. "Modelling the Choice of Residential Location." Cowles Foundation for Research in Economics, Yale University 477, Cowles Foundation for Research in Economics, Yale University.

McFadden, Daniel, Clifford Winston, and Axel Boersch-Supan. 1986. "Joint estimation of freight transportation decisions under nonrandom sampling." *Analytical Studies in Transport Economics*, , ed. Andrew F.Editor Daughety, 137–158. Cambridge University Press.

Melo, Emerson. 2012. "A representative consumer theorem for discrete choice models in networked markets." *Econ. Lett.*, 117(3): 862–865.

Michaels, Guy. 2008. "The effect of trade on the demand for skill: Evidence from the interstate highway system." *The Review of Economics and Statistics*, 90(4): 683–701.

Mundaca, Gabriela, Jon Strand, and Ian R Young. 2021. "Carbon pricing of international transport fuels: Impacts on carbon emissions and trade activity." *Journal of Environmental Economics and Management*, 110: 102517.

Notteboom, Theo. 2018. "Trade and transport modes: the search for global connectivity through transport networks." In *Handbook of international trade and transportation*. 541–565. Edward Elgar Publishing.

Office of the State Auditor. 2019. "Capitalization and Depreciation of Infrastructure." Society for Economic Dynamics 2019 Meeting Papers 212.

Oum, Tae Hoon. 1979. "A cross sectional study of freight transport demand and rail-truck competition in Canada." *The Bell Journal of Economics*, 463–482.

Oum, Tae Hoon. 1989. "Alternative demand models and their elasticity estimates." *Journal of Transport Economics and policy*, 163–187.

Oyama, Yuki, Yusuke Hara, and Takashi Akamatsu. 2022. "Markovian traffic equilibrium assignment based on network generalized extreme value model." *Trans. Res. Part B: Methodol.*, 135–159.

Papola, Andrea, and Vittorio Marzano. 2013. "A Network Generalized Extreme Value Model for Route Choice Allowing Implicit Route Enumeration." *Comput.-aided civ. infrastruct. eng.*, 28(8): 560–580.

Ranieri, Luigi, Salvatore Digiesi, Bartolomeo Silvestri, and Michele Roccotelli. 2018. "A review of last mile logistics innovations in an externalities cost reduction vision." *Sustainability*, 10(3): 782.

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

Redding, Stephen J. 2020. "Trade and Geography." NBER Working Paper 27821.

Redding, Stephen J, and Matthew A Turner. 2015. "Transportation costs and the spatial organization of economic activity." *Handbook of regional and urban economics*, 5: 1339–1398.

Rich, J., O. Kveiborg, and C.O. Hansen. 2011. "On structural inelasticity of modal substitution in freight transport." *Journal of Transport Geography*, 19(1): 134–146.

Rodrigue, Jean-Paul. 2020. *The geography of transport systems.* Routledge.

Rojanasakul, Mira. 2024. "Panama Canal Drought Slows Cargo Traffic."

Santamaria, Marta. 2020. "Reshaping Infrastructure: Evidence from the division of Germany." *The Warwick Economics Research Paper Series (TWERPS)*.

Shapiro, Joseph S. 2016. "Trade Costs, CO2, and the Environment." *American Economic Journal: Economic Policy*, 8(4): 220–54.

Surface Transportation Board. 2009. "A Study of Competition in the US Freight Railroad Industry and Analysis of Proposals that might Enhance Competition." Report prepared by Laurits R. Christensen Associates.

UN. 2021. "Climate Change Fact Sheet." United Nations Sustainable Transport Conference.

USITC. 1991. "The Economic Effects of Significant US Import Restraints." *Investigation No.332-325, Publication 3201*.

USITC. 1995. "The Economic Effects of Significant US Import Restraints."

Winston, Clifford. 1981. "A Disaggregate Model of the Demand for Intercity Freight Transportation." *Econometrica*, 49(4): 981–1006.

Wong, Woan Foong. 2022. "The round trip effect: Endogenous transport costs and international trade." *American Economic Journal: Applied Economics*, 14(4): 127–66.

Zgonc, Borut, Metka Tekavčič, and Marko Jakšič. 2019. "The impact of distance on mode choice in freight transport." *European Transport Research Review*, 11(1): 1–18.

Part

Online Appendix

Table of Contents

A Theoretical Derivations

This appendix presents derivations for the theory section. Table A.1 lists the notations that are used in the theory framework.

Variable Name	Variable
Multi-layered graph	$\mathscr{G} \equiv (\mathscr{N}, \mathscr{L})$
Successor nodes	$\mathscr{B}(i)$
Predecessor nodes	$\mathscr{F}(i)$
Neighboring nodes	$\mathscr{N}(i)$
Route of length K	$r \equiv \{i = r_0, r_1, \ldots, r_K = j\}$
Elasticity of substitution (goods)	σ
Dispersion parameter (route)	θ
Dispersion parameter (mode)	η
Expected minimal transport cost between i and j	τ_{ij}
(Expected minimal) Edge-specific transport cost between i and k	t_{ik}
Mode-specific transport cost between i and k for mode m	$\tilde{t}_{ik,m}$
Link intensity	π_{ij}^{kl}
Link-sourcing probability	$\pi_{ij,k}$
Link choice probability conditional on sourcing from i to j	π_{ij}^{k}
Mode choice probability conditional on link ik	$\pi_{ij,k}^{m}$
Mode-link-sourcing probability	$\pi_{ij,k,m}$

Table. A.1. Overview of notation

A.1 Derivations for Section 3.2: Recursive Routing Choice

A consumer resides in location j and makes a route-sourcing choice comparing prices across all producer locations i and transportation costs across multiple routes r. We make the assumption that transport cost are multiplicative and that they are furthermore the product of edge-specific transport costsalong the route r, i.e. $\tau_{ij,r}^{-\theta_j} = \left(\prod_{l=1}^{K} t_{r_{t-1},r_l}^{-\theta_j}\right)$. Finally, we assume that the consumer makes a recursive route-sourcing choice, beginning at the destination location j by comparing the sourcing prices across the set of (predecessor) neighboring nodes $k \in \mathscr{B}(j)$ net of the transport cost of traversing via the set of (predecessor) neighboring nodes, t_{kj}. For now - without loss of generality - we simply assume that the consumer faces a routing problem along a single-layered graph. As such, the edge-specific transport cost t_{kj} can be interpreted as the expected minimum link cost of traversing to the neighboring node along a multimodal transport network. In summary, the consumer at location j faces the following set of (recursively defined) prices:

$$p_{ij,k}(\nu) = \frac{t_{kj}\tau_{ik}w_i}{\varepsilon_{ij,k}(\nu)}$$

where $\varepsilon_{ij,k}(\nu)$ is a random variable drawn from a Frechet distribution with cumulative distribution given by,

$$F_{ijk}(\epsilon) = e^{-T_{ijk}\epsilon^{-\theta_j}}$$

The consumer in location j is presented with a set of route-source specific prices across producer locations i and chooses

a route by traversing towards neighboring node k, i.e.

$$\begin{aligned} G_{ijk}(p) &= \Pr\left(P_{ijk} \leq p\right) = 1 - F_{ijk}\left(\frac{t_{kj}\tau_{ik}w_i}{p}\right) \\ &= 1 - e^{-T_{ijk}\left(t_{kj}\tau_{ik}w_i\right)^{-\theta_j}p^{\theta_j}} \end{aligned}$$

We can use this distribution to characterize the lowest route-source specific price that consumer in location j faces. To do so, fix an arbitrary threshold price p. The lowest price will be less than p, unless each route-source specific price is greater than p. We therefore seek to characterize, $G_{ijk}(p) = \Pr\left(P_{ijk} \leq p\right)$, which is given by,

$$\begin{aligned} G_j(p) &= 1 - \prod_{i,k}\left(1 - G_{ijk}(p)\right) \\ &= 1 - \prod_{i,k} e^{-T_{ijk}\left(t_{kj}\tau_{ik}w_i\right)^{-\theta_j}p^{\theta_j}} \\ &= 1 - e^{-\Phi_j p^{\theta_j}} \end{aligned}$$

where $\Phi_j = \sum_{ik} T_{ijk}\left(t_{kj}\tau_{ik}w_i\right)^{-\theta_j}$. If $p_{ij,k}(\nu) = p$ then the probability that ijk is the cost minimizing route-source choice is given by,

$$\begin{aligned} \prod_{i'\neq i,k'\neq k} \Pr\left[P_{ijrm} \geq p\right] &= \prod_{i'\neq i,k'\neq k}\left[1 - G_{i'jk'}\right] \\ &= \prod_{i'\neq i,k'\neq k} e^{-T_{ijk}\left(t_{kj}\tau_{ik}w_i\right)^{-\theta_j}p^{\theta_j}} \\ &= e^{-\left(\sum_{i,k} T_{ijr}\left(t_{kj}\tau_{ik}w_i\right)^{-\theta_j}\right)p^{\theta_j}} \end{aligned}$$

Integrating over all possible equilibrium prices p we can characterize the probability that node k and producer location i are the cost minimizing route-source choices:

$$\begin{aligned} \pi_{ij,k} &= \int_0^\infty \prod_{i'\neq i,k'\neq k}\left[1 - G_{i'jk'}\right] dG_{ijk}(p) \\ &= \int_0^\infty \prod_{i'\neq i,k'\neq k} e^{-T_{i'jk'}\left(t_{k'j'}\tau_{i'k'}w_{i'}\right)^{-\theta_j}p^{\theta_j}} dG_{ijk}(p) \end{aligned}$$

Replacing with $dG_{ijk}(p) = \left[T_{ijk}\left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j}\theta_j p^{\theta_j - 1}\right] e^{-T_{ijk}\left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j}p^{\theta_j}} dp$, we obtain,

$$\begin{aligned} \pi_{ij,k} &= \int_0^\infty \prod_{i'\neq i,k'\neq k} e^{-T_{i'jk'}\left(t_{k'j}\tau_{i'j}w_{i'}\right)^{-\theta_j}p^{\theta_j}} dG_{ijk}(p) \\ &= \int_0^\infty \prod_{i'\neq i,k'\neq k} e^{-T_{i'jk'}\left(t_{k'j}\tau_{i'j}w_{i'}\right)^{-\theta_j}p^{\theta_j}} \left[T_{ijk}\left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j}\theta_j p^{\theta_j - 1}\right] e^{-T_{ijk}\left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j}p^{\theta_j}} dp \\ &= T_{ijk}\left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j} \int_0^\infty \prod_{i,k} e^{-T_{ijk}\left(t_{jk}\tau_{ji}w_i\right)^{-\theta_j}p^{\theta_j}} \left[\theta_j p^{\theta_j - 1}\right] dp \\ &= T_{ijk}\left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j} \int_0^\infty e^{-\left(\sum_{i,k} T_{ijk}\left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j}\right)p^{\theta_j}} \left[\theta_j p^{\theta_j - 1}\right] dp \\ &= T_{ijk}\left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j} \int_0^\infty e^{-\Phi_j p^{\theta_j}} \left[\theta_j p^{\theta_j - 1}\right] dp \\ &= T_{ijk}\left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j} \left[\frac{1}{\Phi_j} e^{-\Phi_j p^{\theta_j}}\right]_0^\infty \\ &= \frac{T_{ijk}\left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j}}{\Phi_j} \end{aligned}$$

Replacing with $\Phi_j = \sum_{ik} T_{ijk} \left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j}$, we obtain,

$$\pi_{ij,k} = \frac{T_{ijk}\left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j}}{\sum_{ik} T_{ijk}\left(t_{kj}\tau_{ij}w_i\right)^{-\theta_j}}$$

replacing $T_{ijk} \equiv \left(\frac{1}{A_i}\right)^{\theta_j}$, we obtain,

$$\pi_{ij,k} = \frac{(w_i/A_i)^{-\theta_j}\left(t_{kj}\tau_{ik}\right)^{-\theta_j}}{\sum_{i\in\mathcal{N}}(w_i/A_i)^{-\theta_j}\sum_{k'\in\mathcal{F}(i)}\left(t_{k'j}\tau_{ik'}\right)^{-\theta_j}}$$

as stated above. Furthermore, the expected minimal price is given by,

$$\begin{aligned} p_{ij,k} &= \mathbb{E}\left[\min_{k\in\mathcal{F}(i)}\left\{\frac{t_{kj}\tau_{ik}w_i}{\varepsilon_{ij,k}(v)}\right\}\right] \\ &\propto \sum_{i\in\mathcal{N}}(w_i/A_i)^{-\theta_j}\sum_{k'\in\mathcal{B}(i)}\left(t_{k'j}\tau_{ik'}\right)^{-\theta_j} \\ &= \sum_{i\in\mathcal{N}}(w_i/A_i)^{-\theta_j}\,\tau_{ij}^{-\theta_j} \end{aligned}$$

where in the last line we have used the result that if we consider separately the routing problem this implies an expected transport cost that is in itself recursively defined (see Equation (1), Section 3.2),

$$\tau_{ij} = \mathbb{E}\left[\min_{k\in\mathcal{F}(i)}\left\{\frac{t_{ik}\tau_{kj}}{\varepsilon_{kj}}\right\}\right] \propto \left(\sum_{k\in\mathcal{B}(i)}\left(t_{ik}\tau_{kj}\right)^{-\theta_j}\right)^{-\frac{1}{\theta_j}}$$

which expressed the transport cost as an index of the continuation values along the different edges of the graph.

A.2 Derivations for Section 3.3: Combined Route and Sourcing Choice, and Gravity

We follow the derivations for the general equilibrium gravity equation in Anderson and van Wincoop (2003) to derive a gravity formula, but allowing for generalized substitution patterns. Let us start by examining the expenditure shares, which are given by,

$$\pi_{ij} = \frac{\delta_i^{-\theta_j}\tau_{ij}^{-\theta_j}}{\sum_{i\in N}\delta_i^{-\theta_j}\tau_{ij}^{-\theta_j}} = \frac{\delta_i^{-\theta_j}\tau_{ij}^{-\theta_j}}{P_j^{\theta_j}}$$

where we have posited the existence of δ_i which is a generic origin-shifting gravity constant that is to be solved for. Combining this sourcing share expression to express bilateral trade flows, we obtain,

$$X_{ij} = \left(\frac{\delta_i\tau_{ji}}{P_j}\right)^{-\theta_j} Y_j$$

where we define the price index, allowing for node-specific substitution elasticities θ_j, i.e.

$$P_j = \left[\sum_{i\in N}\delta_i^{-\theta_j}\tau_{ij}^{-\theta_j}\right]^{-\frac{1}{\theta_j}}$$

We now impose market clearing to derive the expression for the gravity constant, i.e.

$$\begin{aligned}
Y_i &= \sum_j X_{ij} \\
&= \sum_j \left(\frac{\delta_i \tau_{ij}}{P_j} \right)^{-\theta_j} Y_j \\
&= (\delta_i)^{-\theta_i} \sum_j \delta_i^{\left(\theta_i - \theta_j\right)} P_j^{\theta_j} \tau_{ij}^{-\theta_j} Y_j \\
&= (\delta_i)^{-\theta_i} \sum_j \delta_i^{\Delta\theta_j} P_j^{\theta_j} \tau_{ij}^{-\theta_j} Y_j
\end{aligned}$$

where we defined the difference between the substitution parameters between nodes i and j as $\Delta\theta_j \equiv \theta_i - \theta_j$. Solving for δ_i,

$$\delta_i = \left(\frac{Y_i}{\sum_j \delta_i^{\Delta\theta_j} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} Y_j} \right)^{-\frac{1}{\theta_i}}$$

Plugging δ_i back into the gravity equation, we obtain

$$\begin{aligned}
X_{ij} &= \left(\frac{\delta_i \tau_{ij}}{P_j} \right)^{-\theta_j} Y_j = (\delta_i)^{-\theta_j} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} Y_j \\
&= \left(\frac{Y_i}{\sum_j \delta_i^{\Delta\theta_j} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} Y_j} \right)^{\frac{\theta_j}{\theta_i}} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} Y_j \\
&= \left(\frac{Y_i Y_j^{\frac{\theta_i}{\theta_j}}}{y^W} \right)^{\frac{\theta_j}{\theta_i}} \left(\frac{1}{\sum_j \delta_i^{\Delta\theta_j} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} \frac{Y_j}{y^W}} \right)^{\frac{\theta_j}{\theta_i}} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} \\
&= \left(\frac{Y_i Y_j^{\frac{\theta_i}{\theta_j}}}{y^W} \right)^{\frac{\theta_j}{\theta_i}} \left(\frac{1}{\left[\sum_j \delta_i^{\Delta\theta_j} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} \frac{Y_j}{y^W} \right]^{-\frac{1}{\theta_j}}} \right)^{-\theta_j \frac{\theta_j}{\theta_i}} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} \\
&= \left(\frac{Y_i Y_j^{\frac{\theta_i}{\theta_j}}}{y^W} \right)^{\frac{\theta_j}{\theta_i}} \left(\frac{1}{\left[\sum_j \delta_i^{\Delta\theta_j} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} \frac{Y_j}{y^W} \right]^{-\frac{1}{\theta_j}}} \right)^{-\theta_j \frac{\theta_j}{\theta_i}} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} \\
&= \left(\frac{Y_i Y_j^{\frac{\theta_i}{\theta_j}}}{y^W} \right)^{\frac{\theta_j}{\theta_i}} \left(\frac{\left(\tau_{ij}\right)^{\frac{\theta_i}{\theta_j}}}{\left[\sum_j \delta_i^{\Delta\theta_j} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} \frac{Y_j}{y^W} \right]^{-\frac{1}{\theta_j}} P_j^{\frac{\theta_i}{\theta_j}}} \right)^{-\theta_j \frac{\theta_j}{\theta_i}} \\
&= \left(\frac{Y_i Y_j^{\frac{\theta_i}{\theta_j}}}{y^W} \right)^{\frac{\theta_j}{\theta_i}} \left(\frac{\left(\tau_{ij}\right)^{\frac{\theta_i}{\theta_j}}}{\Pi_i P_j^{\frac{\theta_i}{\theta_j}}} \right)^{-\theta_j \frac{\theta_j}{\theta_i}}
\end{aligned}$$

where the producer index is given by,

$$\Pi_i = \left(\sum_j \delta_i^{\Delta\theta_j} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} \frac{Y_j}{y^W} \right)^{-\frac{1}{\theta_j}}$$

We can substitute in the equilibrium scaled prices to get,

$$\begin{aligned}
P_j &= \left[\sum_i (\delta_i)^{-\theta_j} \left(\tau_{ij} \right)^{-\theta_j} \right]^{-\frac{1}{\theta_j}} \\
&= \left[\sum_i \left(\frac{Y_i}{\sum_j \delta_i^{\Delta\theta_j} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} y_j} \right)^{\frac{\theta_j}{\theta_i}} \left(\tau_{ij} \right)^{-\theta_j} \right]^{-\frac{1}{\theta_j}} \\
&= \left[\sum_i \left(\frac{\frac{Y_i}{y^W}}{\sum_j \delta_i^{\Delta\theta_j} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} \frac{Y_j}{y^W}} \right)^{\frac{\theta_j}{\theta_i}} \left(\tau_{ij} \right)^{-\theta_j} \right]^{-\frac{1}{\theta_j}} \\
&= \left[\sum_i \left(\frac{Y_i}{y^W} \right)^{\frac{\theta_j}{\theta_i}} \left(\frac{\left(\tau_{ij} \right)^{\frac{\theta_i}{\theta_j}}}{\left(\sum_j \delta_i^{\Delta\theta_j} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} \frac{Y_j}{y^W} \right)^{-\frac{1}{\theta_j}}} \right)^{-\theta_j \frac{\theta_j}{\theta_i}} \right]^{-\frac{1}{\theta_j}} \\
&= \left[\sum_i \tau_{ij}^{-\theta_j} \left(\frac{\frac{Y_i}{y^W}}{\Pi_i} \right)^{\frac{\theta_j}{\theta_i}} \right]^{-\frac{1}{\theta_j}}
\end{aligned}$$

In conclusion we have the following producer and consumer price indices respectively,

$$\Pi_i \equiv \left(\sum_j \delta_i^{\Delta\theta_j} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} \gamma_j \right)^{-\frac{1}{\theta_j}}$$

$$P_j = \left[\sum_i \tau_{ij}^{-\theta_j} \left(\frac{\gamma_i}{\Pi_i} \right)^{\frac{\theta_j}{\theta_i}} \right]^{-\frac{1}{\theta_j}}$$

and we obtain the following gravity equation,

$$X_{ij} = \tau_{ij}^{-\theta_j} \times \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}} \times \frac{\delta_j}{P_j^{-\theta_j}} \tag{24}$$

where $\delta_i = Y_i$ and $\gamma_i = \frac{Y_j}{y^w}$ represent the gravity constants, P_j is the consumer price index or inward market access term, and Π_i is the producer price index or outward market access term.[1] For symmetric substitution parameters (if $\theta_j = \theta_i = \theta$), both the gravity formula and price indices collapse to their traditional form.

[1] Producer price index is $\Pi_i \equiv \left(\sum_j \delta_i^{(\theta_i - \theta_j)} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta_j} \gamma_j \right)^{-\frac{1}{\theta_j}}$ while consumer price index is $P_j = \left[\sum_i \tau_{ij}^{-\theta_j} \left(\frac{\gamma_i}{\Pi_i} \right)^{\frac{\theta_j}{\theta_i}} \right]^{-\frac{1}{\theta_j}}$.

A.3 Derivations for Section 3.3: Combined Routing and Sourcing Choice, and Recursive Equilibrium

This section provides additional derivations for the equilibrium equations in Section 3.3.

In terms of market clearing and trade balance we have the following two equilibrium conditions,

$$Y_i = \sum_j X_{ij}$$
$$E_i = \sum_j X_{ji}$$

Starting with the first equilibrium condition we can derive the first part of Equation (4):

$$\gamma_i = \sum_j X_{ij}$$
$$\gamma_i = \sum_j \tau_{ij}^{-\theta_j} \times \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}} \times \frac{\delta_j}{P_j^{-\theta_j}}$$
$$\Pi_i^{-\theta_i} = \sum_j \tau_{ij}^{-\theta_j} \times \Pi_i^{\theta_j \frac{\theta_j}{\theta_i} - \theta_i} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \times \frac{\delta_j}{P_j^{-\theta_j}}$$
$$\Pi_i^{-\theta_i} = \sum_j \tau_{ij}^{-\theta_j} \times \Pi_i^{\Delta_{i,j}} \times (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \times \frac{\delta_j}{P_j^{-\theta_j}}$$
$$\Pi_i^{-\theta_i} = \tau_{ii}^{-\theta_i} \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{j \neq i} \tau_{ij}^{-\theta_j} \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \frac{\delta_j}{P_j^{-\theta_j}}$$
$$\Pi_i^{-\theta_i} = \tau_{ii}^{-\theta_i} \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{j \neq i} \left(\sum_{k \in \mathscr{B}(k)} t_{ik}^{-\theta_j} \tau_{kj}^{-\theta_j} \right) \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \frac{\delta_j}{P_j^{-\theta_j}}$$
$$\Pi_i^{-\theta_i} = \tau_{ii}^{-\theta_i} \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{j \neq i} \left(\sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta_j} \tau_{kj}^{-\theta_j} \right) \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \frac{\delta_j}{P_j^{-\theta_j}}$$
$$\Pi_i^{-\theta_i} = \tau_{ii}^{-\theta_i} \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta_i} \left(-\tau_{ki}^{-\theta_i} \frac{\delta_i}{p_i^{-\theta_i}} + \sum_j t_{ik}^{\Delta\theta_j} \tau_{kj}^{-\theta_j} \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \frac{\delta_j}{P_j^{-\theta_j}} \right)$$
$$\Pi_i^{-\theta_i} = \left(t_{ii}^{-\theta_i} + \sum_{k \in \mathscr{F}(i)} (t_{ik} \tau_{ki})^{-\theta_i} \right) \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta_i} \left(-\tau_{ki}^{-\theta_i} \frac{\delta_i}{p_i^{-\theta_i}} + \sum_j t_{ik}^{\Delta\theta_j} \tau_{kj}^{-\theta_j} \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \frac{\delta_j}{P_j^{-\theta_j}} \right)$$
$$\Pi_i^{-\theta_i} = \left(t_{ii}^{-\theta_i} \right) \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta_i} \left(\sum_j t_{ik}^{\Delta\theta_j} \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \tau_{kj}^{-\theta_j} \frac{\delta_j}{P_j^{-\theta_j}} \right)$$
$$\Pi_i^{-\theta_i} = t_{ii}^{-\theta_i} \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta_i} \tilde{\Pi}_k^{-\theta_i}$$

where in the second line we defined, $\Pi_i^{\Delta_{i,j}} \equiv \Pi_i^{\theta_j \frac{\theta_j}{\theta_i} - \theta_i}$ and where $\tilde{\Pi}_k^{-\theta_i} = \left(\sum_j t_{ik}^{\Delta\theta_j} \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \tau_{kj}^{-\theta_j} \frac{\delta_j}{P_j^{-\theta_j}} \right)^{\frac{\theta_i}{\theta_j}}$ is the approximate price index that controls for differences in elasticities of substitution between neighboring nodes. And where in the fourth line we have used the fact that at the penultimate node the set of predecessor nodes of k ($\mathscr{B}(k)$) coincides with set of successor nodes of i ($\mathscr{F}(i)$). And finally, where in the last line we used the definition of the recursive transport cost (Equation (1)), $\tau_{ii}^{-\theta} = \left(t_{ii}^{-\theta} + \sum_{k \in \mathscr{F}(i)} \left(t_{ik} \tau_{kj} \right)^{-\theta} \right)$.

Continuing with the second equilibrium condition we can derive the second part of Equation (4),

$$\delta_j = \sum_i X_{ij}$$

$$\delta_j = \sum_i \tau_{ij}^{-\theta_j} \times \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}} \times \frac{\delta_j}{P_j^{-\theta_j}}$$

$$P_j^{-\theta_j} = \sum_i \tau_{ij}^{-\theta_j} \times \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}}$$

$$P_j^{-\theta_j} = \tau_{jj}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{i \neq j} \tau_{ij}^{-\theta_j} \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}}$$

$$P_j^{-\theta_j} = \tau_{jj}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{i \neq j} \left(\sum_{k \in \mathscr{B}(j)} \tau_{ik}^{-\theta_j} t_{kj}^{-\theta_j} \right) \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}}$$

$$P_j^{-\theta_j} = \tau_{jj}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta}} + \sum_{k \in \mathscr{B}(j)} t_{kj}^{-\theta_j} \left(-\tau_{jk}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_i \tau_{ik}^{-\theta_j} \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}} \right)$$

$$P_j^{-\theta_j} = \left(t_{jj}^{-\theta_j} + \sum_{k \in \mathscr{B}(j)} \left(t_{jk} \tau_{kj} \right)^{-\theta_j} \right) \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{k \in \mathscr{B}(j)} t_{kj}^{-\theta_j} \left(-\tau_{jk}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_i \tau_{ik}^{-\theta_j} \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}} \right)$$

$$P_j^{-\theta_j} = \left(t_{jj}^{-\theta_j} \right) \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{k \in \mathscr{B}(j)} t_{kj}^{-\theta_j} \left(\sum_i \tau_{ik}^{-\theta_j} \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}} \right)$$

$$P_j^{-\theta_j} = t_{jj}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{k \in \mathscr{B}(j)} t_{kj}^{-\theta_j} P_k^{-\theta_j}$$

where in the last line we used the definition of the recursive transport cost (Equation (1)), $\tau_{ii}^{-\theta} = \left(t_{ii}^{-\theta} + \sum_{k \in \mathscr{F}(i)} \left(t_{ik} \tau_{kj} \right)^{-\theta} \right)$. Therefore we have,

$$P_j^{-\theta_j} = t_{jj}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{k \in \mathscr{B}(j)} t_{kj}^{-\theta_j} P_k^{-\theta_j}$$

$$\Pi_i^{-\theta_i} = \left(t_{ii}^{-\theta_i} \right) \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta_i} \left(\sum_j t_{ik}^{\Delta \theta_j} \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \tau_{kj}^{-\theta_j} \frac{\delta_j}{P_j^{-\theta_j}} \right)$$

Notice that for the case where we have $\theta_i = \theta_j$, this reduces to,

$$\Pi_i^{-\theta} = \left(t_{ii}^{-\theta} \right) \frac{\delta_i}{P_i^{-\theta}} + \sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta} \left(\sum_j \tau_{kj}^{-\theta} \frac{\delta_j}{P_j^{-\theta}} \right) = \left(t_{ii}^{-\theta} \right) \frac{\delta_i}{P_i^{-\theta}} + \sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta} \Pi_k^{-\theta}$$

$$P_j^{-\theta} = t_{jj}^{-\theta} \frac{\gamma_j}{\Pi_j^{-\theta}} + \sum_{k \in \mathscr{B}(j)} t_{kj}^{-\theta} P_k^{-\theta}$$

We can further express the price indices in terms of equilibrium labor and income allocations across space. We assume welfare equalization, $W_j = \frac{w_j}{P_j} u_j$, and assume localized productivity (A_i) and amenity spillovers (u_i) that depend on the density of workers in a locality, $A_i = \bar{A}_i L_i^{\alpha}$ and $u_i = \bar{u}_i L_i^{\beta}$, where $\bar{A}_i$ is the exogenous productivity component at location i and α determines the extent to which productivity is affected by the local population L_i (productivity spillovers), and where $\bar{u}_i$ is the exogenous utility derived from living in location i and β governs the extent to which amenities are affected by the location population (amenity spillovers). We rewrite the producer and consumer price index respectively,

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

$$\Pi_i = A_i L_i Y_i^{-\frac{\theta+1}{\theta}} \iff$$
$$\Pi_i = A_i L_i^{\alpha+1} Y_i^{-\frac{\theta+1}{\theta}}$$

As a result, we have $P_j = \frac{1}{W} \bar{u}_j L_j^{\beta-1} Y_j$ and $\Pi_i = \bar{A}_i L_i^{1+\alpha} Y_i^{-\frac{\theta_i+1}{\theta_i}}$.[2] Specifically, given the productivity and amenity fundamentals $\left\{\bar{A}_i, \bar{u}_i\right\}$ as well as the aggregate transport costs, the system of $2N$ equations can be solved for the $2N$ endogenous equilibrium values. The equilibrium system determines the endogenous variables via the interaction of price indices along the network topology, where transport cost is endogenously determined as part of the equilibrium system.

A.4 Derivations for Section 3.4: Nested Mode Choice and Aggregate Transport Cost

As presented in the previous section, we consider a consumer that resides in location j and makes a route-sourcing choice by choosing sequentially edges along the graph. This provides a convenient characterization of routing that avoids the curse of dimensionality by expressing the problem as a recursive problem instead of considering the universe of possible routes along a possibly high-dimensional graph. To furthermore accommodate multimodal routing choices, we incorporate a nested modal choice. In that setting, conditional on the neighboring node chosen, the consumer makes a modal choice by choosing the cost minimizing mode out of all modes along the edge, i. to traverse the edge (subject to an extreme value distributed cost shock). To fix ideas, consider a consumer in location j having chosen to route towards the neighboring node k. The consumer then compares all the different modes that are available along this edge, i.e. $m \in \mathscr{M}_{jk}$, where $\mathscr{M}_{jk}$ is the set of modes available between nodes $j$and k, and where the edge mode specific transport costs given by,

$$\tilde{t}_{jk,m} = \begin{cases} t_{jk,m} & \text{if} \quad m = 1 \\ s_{jj,m} \tau_{jk,m} s_{kk,m} & \text{if} \quad m \neq 1 \end{cases}$$

where for the primary mode no switching cost is required, but for any non-primary mode ($m \neq 1$) a switching cost ist imposed, while $\tau_{jk,m}$ refers to the iceberg transport cost of traversing the edge between node j and k along mode m. Notice that this specification is extremeley general and allows for geographies where non-primary modes might connect an entirely different set of nodes than primary nodes. The consumer then faces a cost minimizing choice subject to an extreme-value distributed cost shock, i.e.

$$\min_{m \in \mathscr{M}(j,k)} \left\{ \frac{t_{jk}}{\varepsilon_m} \right\}$$

where the properties of the Frechet distribution implies that the expected cost minimizing transport along any mode between nodes j and k is given by,

$$t_{ik}^{-\theta} \propto \left(\sum_{m \in \mathscr{M}(i,k)} \tilde{t}_{ik,m}^{-\eta} \right)^{\frac{\theta}{\eta}|}$$

Overall, the consumer's route-sourcing choice can be written as a nested minimization problem where we can characterize

[2] We can furthermore rewrite the equations in terms of rescaled variables, $\{y_i, l_i\}$, where we define shares of world income in location i, $y_i \equiv \frac{Y_i}{Y^W}$, and shares of total labor in location i, $l_i \equiv \frac{L_i}{L^W}$. To illustrate this, we did this explicitly for the counterfactual equilibrium system in the Online Appendix A.8.

the overall expected minimal cost, i.e.

$$\begin{aligned} p_{ij} &= \mathbb{E}\left[\min_{j\in\mathscr{F}(i)}\left\{\mathbb{E}\left[\min_{m\in\mathscr{M}(i,k)}\left\{\frac{t_{ik,m}}{\varepsilon_m}\right\}\right]\tau_{jd}w_i\frac{1}{\varepsilon_{ij}}\right\}\right] \\ &\propto \sum_{i\in\mathscr{N}}(w_i/A_i)^{-\theta}\sum_{k'\in\mathscr{F}(i)}\left(\left(\sum_{m\in\mathscr{M}(i,k)}\tilde{t}_{ik,m}^{-\eta}\right)^{-\frac{1}{\eta}}\tau_{k'i}\right)^{-\theta} \\ &= \sum_{i\in\mathscr{N}}(w_i/A_i)^{-\theta}\,\tau_{ij}^{-\theta} \end{aligned}$$

where in the last line we again have used the definition of the transportation cost in terms of the edge-specific costs and associated continuation values along neighboring nodes.

$$\tau_{id} = \mathbb{E}\left[\min_{j\in\mathscr{F}(i)}\left\{\frac{t_{ij}\tau_{jd}}{\varepsilon_{ij}}\right\}\right] \propto \left(\sum_{j\in\mathscr{F}(i)}\left(\left(\sum_{m\in\mathscr{M}(i,k)}t_{ik,m}^{-\eta}\right)^{-\frac{1}{\eta}|}\tau_{jd}\right)^{-\theta}\right)^{-\frac{1}{\theta}}$$

which expressed the transport cost as an index of the continuation values along the different edges of the graph.

The mode-route-sourcing choice probability is therefore given by,

$$\begin{aligned} \pi_{ij,k,m} &= \frac{(w_i/A_i)^{-\theta}\,(t_{ik}\tau_{ki})^{-\theta}}{\sum_{i\in\mathscr{N}}(w_i/A_i)^{-\theta}\sum_{k'\in\mathscr{F}(i)}(t_{ik'}\tau_{k'i})^{-\theta}}\frac{\tilde{t}_{ik,m}^{-\eta}}{t_{ik}^{-\eta}} \\ &\equiv \pi_{ij,k}\times\pi_{ik}^{m} \end{aligned}$$

where in the last line, we can decompose this probability into the link choice probability ($\pi_{ij,k}$) and the mode choice probability along the link (π_{ik}^m, as discussed in Equation (6)).

A.5 Isomorphism with Allen and Arkolakis (2022)

It might be insightful to link the recursive expression in our paper to the approach in AA2022 that instead relies on explicit enumeration of the universe of paths and utilizes matrix algebra to express the expected minimum transport cost in terms of the leontief inverse of the adjacency matrix that captures the underlying infrastructure network. As we will show, the model is isomorphic in the case where there is a simple unimodal transport network. Furthermore, we will also show that for the simplified case where the mode choice elasticity is equal to the route choice elasticity, we can show that their approach can be extended to provide a clean expression of the multimodal transport cost in terms of (leontief inverses) of the infrastructure matrices. First, let us demonstrate the isomorphism for unimodal transport network. Consider the recursive transport cost stated above,

$$\tau_{id} = \mathbb{E}\left[\min_{j\in\mathscr{F}(i)}\left\{\frac{t_{ij}\tau_{jd}}{\varepsilon_{ij}}\right\}\right] = \left(\sum_{j\in\mathscr{F}(i)}\left(t_{ij}\tau_{jd}\right)^{-\theta}\right)^{-\frac{1}{\theta}}$$

If we assume a finite graph, then we can iteratively substitute and obtain a closed-form expression for the endogenous transport cost, i.e.

$$\begin{aligned} (\tau_{id})^{-\theta} &= t_{d-1d}^{-\theta}\sum_{j\in F(i)}t_{ij}^{-\theta}\sum_{j'\in F(j)}\cdots\sum_{k'\in F(d-1)}t_{k'd-1}^{-\theta} \\ &= \sum_{r\in\mathfrak{R}_{ij}}\left(\prod_{l=1}^{K}t_{r_{l-1},r_l}^{-\theta}\right) \\ &= \sum_{K=0}^{\infty}A_{ij}^{K} \\ &= (\mathbf{I}-\mathbf{A})^{-1}\equiv\mathbf{B} \end{aligned}$$

where in the second line we recognize that the recursive substitution on a finite graph results in a characterization of all possible routes of any length along the network. The resulting expression is identical to expression for endogenous transport cost in AA2022. In the third and fourth line we then employ the same argument as in their paper to show that the implied transport cost can be expressed as the leontief inverse of the underlying infrastructure matrix, where $\mathbf{A} \equiv \left[t_{ij}^{-\theta} \right]$ is an $N \times N$ matrix with (i, j) element $t_{ij}^{-\theta}$ and A_{ij}^{K} is the (ij) element of the matrix A to the matrix power K. This shows that in essence, the two different approaches are isomorphic and capture the same underlying endogenous transport cost, albeit in different ways.

We now turn towards showing that for the special case where the mode choice elasticity is equal to the route choice elasticity, we can furthermore extend the approach in AA2022 to derive a clean decomposition of the multimodal transport cost in terms of a set of underlying mode-specific infrastructure matrices. Let us therefore consider the case where $\eta = \theta$. Notice that in this special case, the edge-specific transport cost is given by,

$$t_{ik}^{-\theta} \propto \sum_{m \in \mathscr{M}(i,k)} \tilde{t}_{ik,m}^{-\theta}$$

where $\tilde{t}_{ik,m}^{-\theta}$ are the mode specific traversal costs between node i and k. As above, we can express recursively substitute the expected transport across modes between nodes, and obtain,

$$(\tau_{id})^{-\theta} = t_{d-1d}^{-\theta} \sum_{j \in F(i)} t_{ij}^{-\theta} \sum_{j' \in F(j)} \dots \sum_{k' \in F(d-1)} t_{k'd-1}^{-\theta}$$

By substituting the edge-specific transport cost in terms of the mode-specific cost, we obtain,

$$\begin{aligned}(\tau_{id})^{-\theta} &= \left(\sum_{m \in \mathscr{M}(d-1,d)} \tilde{t}_{d-1d,m}^{-\theta} \right) \sum_{j \in F(i)} \left(\sum_{m \in \mathscr{M}(i,j)} \tilde{t}_{ij,m}^{-\theta} \right) \sum_{j' \in F(j)} \dots \sum_{k' \in F(d-1)} \left(\sum_{m \in \mathscr{M}(k',d-1)} \tilde{t}_{k'd-1,m}^{-\theta} \right) \\ &= \sum_{r \in \Re_{ij}} \left(\prod_{l=1}^{K} \sum_{m \in \mathscr{M}(r_{l-1}, r_l)} \tilde{t}_{r_{l-1}, r_l, m}^{-\theta} \right)\end{aligned}$$

where the second line is a concise way of summarizing all possible uni and multi-modal paths along the multi-layered network.Without loss of generality, let us consider the case of two modes - one primary mode where flows originate and terminate, and a secondary mode that is accessible subject to some switching cost. Define the $(N_1 + N_2) \times (N_1 + N_2)$ matrix $\mathbf{A} = \left[a_{ij} \equiv t_{ij}^{-\theta} \right]$. Notice that this adjacency matrix forms a block partitioned matrix, i.e.

$$\mathbf{A} = \begin{bmatrix} \mathbf{A_1} & \mathbf{S} \\ \mathbf{S}' & \mathbf{A_2} \end{bmatrix}$$

where $\mathbf{A_1} = \left[a_{ij} \right] = \left[t_{ij}^{-\theta} \right]$ is the adjacency matrix for the primary transportation network, $\mathbf{A_2} = \left[a_{i'j'} \right] = \left[t_{i'j'}^{-\theta} \right]$ is the adjacency matrix for the secondary transportation network, and $\mathbf{S} = \left[s_{ii'}^{-\theta} \right]$ is the diagonal matrix that represents linkages between the primary and secondary transportation network. We can write τ_{ij} from by explicitly summing across all possible routes of all possible lengths. To do so, we sum across all locations that are traveled through all the possible paths as follows:

$$\begin{aligned}\tau_{ij}^{-\theta} &= \sum_{r \in \Re_{ij}} \left(\prod_{l=1}^{K} \sum_{m \in \mathscr{M}(r_{l-1}, r_l)} \tilde{t}_{r_{l-1}, r_l, m}^{-\theta} \right) \\ &= \sum_{K=0}^{\infty} \left(\sum_{k_1=1}^{(N_1+N_2)} \sum_{k_2=1}^{(N_1+N_2)} \dots \sum_{k_{K-1}=1}^{(N_1+N_2)} a_{i,k_1} \times a_{k_1,k_2} \times \dots \times a_{k_{K-2},k_{K-1}} \times a_{k_{K-1}j} \right)\end{aligned}$$

By explicitly recognizing that this sum across all locations through all possible paths can be partitioned into unimodal paths on each transportation network and an arbitrary number of switches between transportation modes, we have,

$$\tau_{ij}^{-\theta} = \sum_{t_1=1}^{N} \sum_{t_2=1}^{N} \dots \sum_{t_S=1}^{N} \left(\left(\sum_{K=0}^{\infty} \mathbf{A}_{1,it_1}^{K} \right) \times s_{t_1 t_1'}^{-\theta} \times \dots \times s_{t_s' t_s}^{-\theta} \left(\sum_{K=0}^{\infty} \mathbf{A}_{1,T_S j}^{K} \right) \right)$$

which in matrix notation can be written as,

$$\tau_{ij}^{-\theta} = \sum_{K=0}^{\infty}\left(\left(\sum_{K=0}^{\infty}\mathbf{A}_1^K\right)\left(\mathbf{S}\left(\sum_{K=0}^{\infty}\mathbf{A}_2^K\right)\mathbf{S}'\right)\right)^K\left(\sum_{K=0}^{\infty}\mathbf{A}_1^K\right)$$

To simplify this expression let us first define the Leontief inverse for each infrastructure matrix separately, i.e.

$$\sum_{K=0}^{\infty}\mathbf{A}_1^K = (\mathbf{I}-\mathbf{A}_1)^{-1} \equiv \mathbf{B}$$

$$\sum_{K=0}^{\infty}\mathbf{A}_2^K = (\mathbf{I}-\mathbf{A}_2)^{-1} \equiv \mathbf{C}$$

We also define - for convenience - the sandwich matrix that adjusts the transport cost along the secondary transportation network for switching costs and therefore traces out the option value of having access to the secondary transportation network,

$$\mathbf{S}\left(\sum_{K=0}^{\infty}\mathbf{A}_2^K\right)\mathbf{S}' \equiv \mathbf{D}$$

From matrix calculus we can restate the following result that relates the inverse of the Schur complement of the partitioned infrastructure matrix to the geometric sum of matrix operations, specifically,

$$\sum_{K=0}^{\infty}\left(\mathbf{B}^{-1}\mathbf{D}\right)^K\mathbf{B}^{-1} = (\mathbf{B}-\mathbf{D})^{-1} \equiv \mathbf{E}$$

Applying this result, we can write,

$$\begin{aligned}\tau_{ij}^{-\theta} &= \sum_{K=0}^{\infty}\left((\mathbf{I}-\mathbf{A}_1)^{-1}\left(\mathbf{S}(\mathbf{I}-\mathbf{A}_2)^{-1}\mathbf{S}'\right)\right)^K(\mathbf{I}-\mathbf{A}_1)^{-1}\\ &= \left[(\mathbf{I}-\mathbf{A}_1)-\mathbf{S}(\mathbf{I}-\mathbf{A}_2)^{-1}\mathbf{S}'\right]^{-1}_{ij}\end{aligned}$$

therefore we can write,

$$\tau_{ij} = e_{ij}^{-\frac{1}{\theta}}$$

Furthermore, the Woodbury matrix identity (see e.g. Horn and Johnson (2012)) states,

$$(\mathbf{A}+\mathbf{UCV})^{-1} = \mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{U}\left(\mathbf{C}^{-1}+\mathbf{VA}^{-1}\mathbf{U}\right)^{-1}\mathbf{VA}^{-1}$$

which implies

$$\begin{aligned}\tau_{ij}^{-\theta} &= \left[(\mathbf{I}-\mathbf{A}_1)-\mathbf{S}(\mathbf{I}-\mathbf{A}_2)^{-1}\mathbf{S}'\right]^{-1}_{ij}\\ &= \left[\mathbf{B}+\mathbf{BS}(\mathbf{A}/\mathbf{A}_1)^{-1}\mathbf{S}'\mathbf{B}\right]_{ij}\\ &= \left[(\mathbf{I}-\mathbf{A}_1)^{-1}+(\mathbf{I}-\mathbf{A}_1)^{-1}\mathbf{S}(\mathbf{A}/\mathbf{A}_1)^{-1}\mathbf{S}'(\mathbf{I}-\mathbf{A}_1)^{-1}\right]_{ij}\end{aligned}$$

where $\mathbf{A}/\mathbf{A}_1 := (\mathbf{I}-\mathbf{A}_1)^{-1}-\mathbf{S}(\mathbf{A}/\mathbf{A}_1)^{-1}\mathbf{S}'$ defines the Schur complement of the adjacency matrix A. The expressions corresponds to the expression given in the main text and intuitively decomposes the transport cost into a component that originates from the unimodal paths and another component that originates from the multimodal paths. This result can also directly be obtained by applying to the partitioned matrix A the formula for the inverse of block-partitioned matrices (see e.g. Horn and Johnson (2012)).

A.6 Derivations for Section 3.5: Congestion and Traffic Flows

We characterize equilibrium traffic at different nodes of the transportation network. First, we utilize the recursive routing choice to characterize aggregate traffic between any two nodes across any mode. Second, we characterize mode-specific traffic between nodes. Finally, we characterize traffic at terminal stations.

A.6.1 Traffic Flows on the Aggregate Network

We start by characterizing the aggregate traffic between any two nodes across any mode. To do so we start by restating the (recursively defined) sourcing and link choice probability which is given by:

$$\pi_{ij,k} = \frac{(w_i/A_i)^{-\theta}\left(t_{jk}\tau_{ki}\right)^{-\theta}}{\sum_{i\in\mathcal{N}}(w_i/A_i)^{-\theta}\sum_{k'\in\mathcal{F}(i)}\left(t_{jk'}\tau_{k'i}\right)^{-\theta}}$$

which can be decomposed in the sourcing share and the link choice probability conditional on the sourcing choice, i.e.

$$\begin{aligned}
\pi_{ij,kl} &= \frac{(t_{ik}\tau_{ki})^{-\theta}}{\sum_{k'\in\mathcal{F}(i)}(t_{ik'}\tau_{k'i})^{-\theta}}\frac{\tau_{ij}^{-\theta}p_i^{-\theta}}{\sum_{i\in\mathcal{N}}\tau_{ij}^{-\theta}p_i^{-\theta}} \\
&= \frac{(t_{ik}\tau_{ki})^{-\theta}}{\tau_{ij}^{-\theta}}\frac{\tau_{ij}^{-\theta}p_i^{-\theta}}{\sum_{i\in\mathcal{N}}\tau_{ij}^{-\theta}p_i^{-\theta}} \\
&= \pi_{ij}^{k}\times\pi_{ij}
\end{aligned}$$

The previous derivations only characterize the probability of choosing neighboring link k when routing between i and j. Notice that we can construct the probability of using any edge kl when transporting goods from i to j in straightforward way,

$$\begin{aligned}
\pi_{kj}^{kl} &= \frac{t_{kl}^{-\theta}\tau_{lj}^{-\theta}}{\tau_{kj}^{-\theta}} \\
&= \frac{\tau_{ik}^{-\theta}t_{kl}^{-\theta}\tau_{lj}^{-\theta}}{\tau_{ik}^{-\theta}\tau_{kj}^{-\theta}} \\
&= \frac{\tau_{ik}^{-\theta}t_{kl}^{-\theta}\tau_{lj}^{-\theta}}{\tau_{ij}^{-\theta}} \\
&= \left(\frac{\tau_{ij}}{\tau_{ik}t_{kl}\tau_{ij}}\right)^{\theta}
\end{aligned}$$

which gives us Equation (7), Section 3.5. Notice that this arises naturally due to the markovian property of the recursive routing choice. To characterize traffic between nodes k and l along any mode, we characterize the share of goods that are being sourced from any location i to any location j and use the link kl along the way, i.e.

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

where in the last line we used the definition of the consumer and producer market access terms. Furthermore, replacing market access terms,

$$\begin{aligned}
P_i &= \frac{1}{\bar{W}}\bar{u}_i L_i^{\beta-1}Y_i \\
\Pi_i &= \bar{A}_i L_i^{1+\alpha}Y_i^{-\frac{\theta+1}{\theta}}
\end{aligned}$$

we obtain,

$$\begin{aligned}\Xi_{kl} &= t_{kl}^{-\theta} \times \left(\frac{1}{\bar{W}}\bar{u}_k L_k^{\beta-1} Y_k\right)^{-\theta} \times \left(\bar{A}_l L_l^{1+\alpha} Y_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta} \\ &= t_{kl}^{-\theta}\left(\frac{\bar{L}^{-(\alpha+\beta)\theta}}{W^{-\theta}}\right)\bar{L}\bar{A}_l^{-\theta}\bar{u}_k^{-\theta} l_k^{-\theta(\beta-1)} l_l^{-\theta(1+\alpha)} y_k^{-\theta} y_l^{(1+\theta)} \\ &= t_{kl}^{-\theta}\chi\bar{L}\bar{A}_l^{-\theta}\bar{u}_k^{-\theta} l_k^{-\theta(\beta-1)} l_l^{-\theta(1+\alpha)} y_k^{-\theta} y_l^{(1+\theta)}\end{aligned}$$

where in the last line we use the definition $\chi = \frac{\bar{L}^{(\alpha+\beta)\theta}}{W^{\theta}}$. We have,

$$\Xi_{kl} = t_{kl}^{-\theta}\chi^{-1}\bar{L}\bar{A}_l^{-\theta}\bar{u}_k^{-\theta} l_k^{-\theta} l_l^{-\theta} y_k^{-\theta} y_l^{(1+\theta)}$$

which characterizes aggregate flows between neighboring nodes k and l in terms of the endogenous variables along the network.

A.6.2 Traffic Flows on the Mode-Specific Networks

In the next step we now turn towards characterizing the probability of sourcing from location j to location i choosing neighboring node k as the cost-minimizing routing choice and furthermore opting for mode m between node i and k. The nested choice implies that this choice probability is given by,

$$\pi_{ijk,m} = \frac{\tilde{t}_{ik,m}^{-\eta}}{t_{ik}^{-\eta}} \frac{\left(t_{ik}\tau_{kj}\right)^{-\theta} p_i^{-\theta}}{\sum_{i\in\mathcal{N}} \tau_{ij}^{-\theta} p_i^{-\theta}}$$

which can be decomposed in the sourcing share, link choice probability conditional on sourcing choice, and the modal share conditional on both sourcing and link choice, i.e.

$$\begin{aligned}\pi_{ij,kl,m} &= \frac{\tilde{t}_{ik,m}^{-\eta}}{t_{ik}^{-\eta}} \frac{\left(t_{ik}\tau_{ki}\right)^{-\theta}}{\sum_{k'\in\mathcal{F}(i)}\left(t_{ik'}\tau_{k'i}\right)^{-\theta}} \frac{\tau_{ij}^{-\theta} p_i^{-\theta}}{\sum_{i\in\mathcal{N}} \tau_{ij}^{-\theta} p_i^{-\theta}} \\ &= \pi_{ik}^{m} \times \pi_{ij}^{k} \times \pi_{ij}\end{aligned}$$

We can apply similar calculations as above to extend this for any link kl when transporting goods from i to j, i.e.

$$\pi_{ij,kl,m} = \frac{\tilde{t}_{kl,m}^{-\eta}}{t_{kl}^{-\eta}} \frac{\tau_{ik}^{-\theta} t_{kl}^{-\theta} \tau_{lj}^{-\theta}}{\tau_{ij}^{-\theta}} \frac{\tau_{ij}^{-\theta} p_i^{-\theta}}{\sum_{i\in\mathcal{N}} \tau_{ij}^{-\theta} p_i^{-\theta}}$$

Given this choice probability we can characterize the mode-specific traffic between neighboring nodes k and l, i.e.

$$\begin{aligned}\Xi_{kl,m} &= \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} \pi_{ijk,m} E_j \iff \\ &= \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} \pi_{ik}^{m} \times \pi_{ij}^{k} \times X_{ij} \\ &= \frac{\tilde{t}_{kl,m}^{-\eta}}{t_{kl}^{-\eta}} \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} \frac{\tau_{ik}^{-\theta} t_{kl}^{-\theta} \tau_{lj}^{-\theta}}{\tau_{ij}^{-\theta}} \times \tau_{ij}^{-\theta} \frac{Y_i}{\Pi_i^{-\theta}} \frac{E_j}{P_j^{-\theta}} \iff \\ &= \frac{\tilde{t}_{kl,m}^{-\eta}}{t_{kl}^{-\eta}} \times t_{kl}^{-\theta} \sum_{i\in\mathcal{N}} \tau_{ik}^{-\theta} \frac{Y_i}{\Pi_i^{-\theta}} \sum_{j\in\mathcal{N}} \tau_{lj}^{-\theta} \frac{E_j}{P_j^{-\theta}}, \\ &= \frac{\tilde{t}_{kl,m}^{-\eta}}{t_{kl}^{-\eta}} \times t_{kl}^{-\theta} \times P_k^{-\theta} \times \Pi_l^{-\theta} \\ &= \tilde{t}_{kl,m}^{-\eta} \times t_{kl}^{\eta-\theta} \times P_k^{-\theta} \times \Pi_l^{-\theta}\end{aligned}$$

which gives us Equation (9) (Section 3.5), an expression for mode-specific traffic in terms of market access measures and the aggregate and mode specific iceberg transport cost along the edge. Note, that the nested formulation implies that

mode-specific traffic is the conditional mode specific share of aggregate traffic, i.e. $\Xi_{kl,m} = \pi^m_{ik} \times \Xi_{kl}$.

A.6.3 Traffic Flows at Terminals

In the final step we characterize the traffic at mode-specific terminals that allow multimodal movements between nodes kl. To do so we start by characterizing the probability of sourcing from location j to location i choosing any neighboring node $k' \in \mathscr{F}(i)$, but crucially choosing an alternative non-primary mode of transport and therefore traversing through a terminal while routing. This choice probability can be characterized in the following way,

$$\begin{aligned}\pi_{ij,kk,m} &= \sum_{k\in\mathscr{F}(i)} \frac{\tilde{t}^{-\eta}_{ik,m}}{t^{-\eta}_{ik}} \frac{t^{-\theta}_{ik}\tau^{-\theta}_{kj}p^{-\theta}_{j}}{\sum_{j\in 0}\sum_{k'} t_{ik}\tau_{k_j j}p^{-\theta}_{j}} \\ &= s^{-\eta}_{ii} \sum_{k\in\mathscr{F}(i)} \frac{\left(\tau_{ik}s_{kk,m}\right)^{-\eta}}{t^{-\eta}_{ik}} \frac{t^{-\theta}_{ik}\tau^{-\theta}_{kj}p^{-\theta}_{j}}{\sum_{j\in 0}\sum_{k'} t_{ik}\tau_{k_j j}p^{-\theta}_{j}}\end{aligned}$$

which can be decomposed into which can again be decomposed into sourcing shares, link choice shares conditional on sourcing choice, and mode choice conditional on link choice, i.e.

$$\begin{aligned}\pi_{ij,kk,m} &= \sum_{k\in\mathscr{F}(i)} \frac{s^{-\eta}_{ii}\left(\tau_{ik}s_{kk,m}\right)^{-\eta}}{t^{-\eta}_{ik}} \frac{\left(t_{ik}\tau_{ki}\right)^{-\theta}}{\sum_{k'\in\mathscr{F}(i)}\left(t_{ik'}\tau_{k'i}\right)^{-\theta}} \frac{\tau^{-\theta}_{ij}p^{-\theta}_{i}}{\sum_{i\in\mathscr{N}}\tau^{-\theta}_{ij}p^{-\theta}_{i}} \\ &= \pi_{ij}\sum_{k\in\mathscr{F}(i)} \pi^m_{ik} \times \pi^k_{ij}\end{aligned}$$

We can apply similar calculations as above to extend this for any link kl when transporting goods from i to j, i.e.

$$\begin{aligned}\pi_{ij,kk,m} &= \sum_{l\in\mathscr{F}(k)} \frac{s^{-\eta}_{kk}\left(\tau_{kl}s_{ll,m}\right)^{-\eta}}{t^{-\eta}_{kl}} \frac{\tau^{-\theta}_{ik}t^{-\theta}_{kl}\tau^{-\theta}_{lj}}{\tau^{-\theta}_{ij}} \frac{\tau^{-\theta}_{ij}p^{-\theta}_{i}}{\sum_{i\in\mathscr{N}}\tau^{-\theta}_{ij}p^{-\theta}_{i}} \\ &= \pi_{ij}\sum_{l\in\mathscr{F}(k)} \pi^m_{kl} \times \pi^{kl}_{ij}\end{aligned}$$

Given this choice probability we can characterize the mode-specific traffic between neighboring nodes k and l, i.e.

$$\begin{aligned}\Xi_{kk,m} &= \sum_{i\in\mathscr{N}}\sum_{j\in\mathscr{N}} \pi_{ij,kk,m}E_j \iff \\ &= \sum_{i\in\mathscr{N}}\sum_{j\in\mathscr{N}}\left(\sum_{l\in\mathscr{F}(k)} \pi^m_{kl} \times \pi^{kl}_{ij}\right) \times X_{ij} \\ &= \sum_{l\in\mathscr{F}(k)} \frac{s^{-\eta}_{kk}\left(\tau_{kl}s_{ll,m}\right)^{-\eta}}{t^{-\eta}_{kl}} \sum_{i\in\mathscr{N}}\sum_{j\in\mathscr{N}} \frac{\tau^{-\theta}_{ik}t^{-\theta}_{kl}\tau^{-\theta}_{lj}}{\tau^{-\theta}_{ij}} \times \tau^{-\theta}_{ij}\frac{Y_i}{\Pi^{-\theta}_{i}}\frac{E_j}{P^{-\theta}_{j}} \iff \\ &= \sum_{l\in\mathscr{F}(k)} \frac{s^{-\eta}_{kk}\left(\tau_{kl}s_{ll,m}\right)^{-\eta}}{t^{-\eta}_{kl}} \times t^{-\theta}_{kl}\sum_{i\in\mathscr{N}}\tau^{-\theta}_{ik}\frac{Y_i}{\Pi^{-\theta}_{i}}\sum_{j\in\mathscr{N}}\tau^{-\theta}_{lj}\frac{E_j}{P^{-\theta}_{j}}, \\ &= \sum_{l\in\mathscr{F}(k)} \frac{s^{-\eta}_{kk}\left(\tau_{kl}s_{ll,m}\right)^{-\eta}}{t^{-\eta}_{kl}} \times t^{-\theta}_{kl} \times P^{-\theta}_{k} \times \Pi^{-\theta}_{l} \\ &= s^{-\eta}_{kk} \times P^{-\theta}_{k} \times \sum_{l\in\mathscr{F}(k)}\left(\tau_{kl}s_{ll,m}\right)^{-\eta} \times t^{\eta-\theta}_{kl} \times \Pi^{-\theta}_{l}\end{aligned}$$

which gives us Equation (10) in Section 3.5.

A.6.4 Congestion on the Primary Transport Network

Incorporating congestion, for the primary mode of transport ($m = 1$), we have the following relationship as in the first part of in Equation (12) (Section 3.5),

$$
\begin{aligned}
t_{kl,1} &= \bar{t}_{kl,1}\left[\Xi_{ij,1}\right]^{\lambda_1} \iff \\
t_{kl,1} &= \bar{t}_{kl,1}\left[t_{kl,1}^{-\eta} \times t_{kl}^{\eta-\theta} \times P_k^{-\theta} \times \Pi_l^{-\theta}\right]^{\lambda_1} \iff \\
t_{kl,1} &= \bar{t}_{kl,1} \times t_{kl,1}^{-\eta\lambda_1} \times t_{kl}^{\lambda_1(\eta-\theta)} \times P_k^{-\theta\lambda_1} \times \Pi_l^{-\theta\lambda_1} \iff \\
t_{kl,1}^{1+\eta\lambda_1} &= \bar{t}_{kl,1} \times t_{kl}^{\lambda_1(\eta-\theta)} \times P_k^{-\theta\lambda_1} \times \Pi_l^{-\theta\lambda_1} \iff \\
t_{kl,1} &= \bar{t}_{kl,1}^{\frac{1}{1+\eta\lambda_1}} \times t_{kl}^{\frac{\lambda_1(\eta-\theta)}{1+\eta\lambda_1}} \times P_k^{\frac{-\theta\lambda_1}{1+\eta\lambda_1}} \times \Pi_l^{\frac{-\theta\lambda_1}{1+\eta\lambda_1}}
\end{aligned}
$$

Substituting the equation above into to derive the expression for traffic flows as in the second part of in Equation (12) (Section 3.5),

$$
\begin{aligned}
\Xi_{kl,1} &= \left(\bar{t}_{kl,1}^{\frac{1}{1+\theta\lambda_1}} \times t_{kl}^{\frac{\lambda_1(\eta-\theta)}{1+\eta\lambda_1}} \times P_k^{-\frac{\theta\lambda_1}{1+\theta\lambda_1}} \times \Pi_l^{-\frac{\theta\lambda_1}{1+\lambda_1}}\right)^{-\theta} \times P_k^{-\theta} \times \Pi_l^{-\theta} \\
&= \bar{t}_{kl,1}^{-\frac{\theta}{1+\theta\lambda_1}} \times t_{kl}^{\frac{-\theta\lambda_1(\eta-\theta)}{1+\eta\lambda_1}} \times P_k^{\frac{\theta\theta\lambda_1}{1+\theta\lambda_1}-\theta} \times \Pi_l^{\frac{\theta\lambda_1\theta}{1+\theta\lambda_1}-\theta} \\
&= \bar{t}_{kl,1}^{-\frac{\theta}{1+\theta\lambda_1}} \times t_{kl}^{\frac{-\theta\lambda_1(\eta-\theta)}{1+\eta\lambda_1)}} \times P_k^{-\frac{\theta}{1+\theta\lambda_1}} \times \Pi_l^{-\frac{\theta}{1+\theta\lambda_1}}
\end{aligned}
$$

A.6.5 Congestion at Intermodal Terminals

For any secondary mode of transportation ($m \neq 1$), we incorporate congestion at terminal stations that connect the primary and secondary networks. Let us characterize the transport cost net of congestion at any terminal, i.e.

$$
\begin{aligned}
s_{kk,m} &= \bar{s}_{kk,m}\left[\Xi_{kk,m}\right]^{\lambda_m} \\
s_{kk,m} &= \bar{s}_{kk,m}\left[s_{kk,m}^{-\eta} \times P_k^{-\theta} \times \sum_{l\in\mathscr{F}(k)}\left(\tau_{kl}s_{ll,m}\right)^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta}\right]^{\lambda_m} \\
s_{kk,m} &= \bar{s}_{kk,m} \times s_{kk,m}^{-\eta\lambda_m} \times P_k^{-\theta\lambda_m} \times \left(\sum_{l\in\mathscr{F}(k)}\left(\tau_{kl}s_{ll,m}\right)^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta}\right)^{\lambda_m} \\
s_{kk,m}^{1+\eta\lambda_m} &= \bar{s}_{kk,m} \times P_k^{-\theta\lambda_m} \times \left(\sum_{l\in\mathscr{F}(k)}\left(\tau_{kl}s_{ll,m}\right)^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta}\right)^{\lambda_m} \\
s_{kk,m} &= \bar{s}_{kk,m}^{1+n\lambda_m} \times P_k^{\frac{-\theta\lambda_m}{1+\eta\lambda_m}} \times \left(\sum_{l\in\mathscr{F}(k)}\left(\tau_{kl}s_{ll,m}\right)^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta}\right)^{\frac{\lambda_m}{1+\eta\lambda_m}} \\
s_{kk,m} &= \bar{s}_{kk,m}^{\frac{1}{1+\eta\lambda m}} \times P_k^{\frac{-\theta\lambda_m}{1+\eta\lambda_m}} \times \left(\sum_{l\in\mathscr{F}(k)}\left(\tau_{kl}s_{ll,m}\right)^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta}\right)^{\frac{\lambda_m}{1+\eta\lambda_m}}
\end{aligned}
$$

Combining both the switching cost of access the non-primary transport network and the terminal cost of exiting it again, we

obtain Equation (14) in Section 3.5:

$$
\begin{aligned}
t_{kl,m} &= \bar{s}_{kk,m}\tau_{kl,m}\bar{s}_{ll,m}\left[\Xi_{kk,m}\right]^{\lambda_m}\left[\Xi_{ll,m}\right]^{\lambda_m} \\
&= \bar{s}_{kk,m}\tau_{kl,m}\bar{s}_{ll,m}\left[\sum_{l\in\mathscr{F}(k)}\Xi_{kl,m}\right]^{\lambda_m}\left[\sum_{k\in\mathscr{B}(l)}\Xi_{kl,m}\right]^{\lambda_m} \\
&= \bar{s}_{kk,m}\tau_{kl,m}\bar{s}_{ll,m}\left[P_k^{-\theta}\times\sum_{l\in\mathscr{F}(k)}\tilde{t}_{kl,m}^{-\eta}\times t_{kl}^{\eta-\theta}\times\Pi_l^{-\theta}\right]^{\lambda_m}\left[\Pi_l^{-\theta}\times\sum_{k\in\mathscr{B}(l)}\tilde{t}_{kl,m}^{-\eta}\times t_{kl}^{\eta-\theta}\times P_k^{-\theta}\right]^{\lambda_m} \\
&= \bar{s}_{kk,m}\tau_{kl,m}\bar{s}_{ll,m}\left[P_k^{-\theta}\times\Pi_{k,m}^{-\theta}\right]^{\lambda_m}\left[\Pi_l^{-\theta}\times P_{l,m}^{-\theta}\right]^{\lambda_m} \\
&= \bar{s}_{kk,m}\times\tau_{kl,m}\times\bar{s}_{ll,m}\times P_k^{-\theta\lambda_m}\times\Pi_l^{-\theta\lambda_m}\times P_{k,m}^{-\theta\lambda_m}\times\Pi_{l,m}^{-\theta\lambda_m}
\end{aligned}
$$

where $P_{l,m}\equiv\sum_{k\in\mathscr{B}(l)}t_{kl,m}^{-\eta}\times t_{kl}^{\eta-\theta}\times P_k^{-\theta}$ and $\Pi_{k,m}\equiv\sum_{l\in\mathscr{F}(k)}t_{kl,m}^{-\eta}\times t_{kl}^{\eta-\theta}\times\Pi_l^{-\theta}$.

A.7 Derivations for Section 3.6: Substitution and Complementarity within the Multimodal Network

The aim of this section is to characterize the impact of a mode-specific transport cost change, $d\ln t_{kl,m}\neq 0$ on mode-specific traffic

$$
d\ln\Xi_{kl,m} = \underbrace{\epsilon_{t_{kl,m}^m}^{\pi_{kl}^m}\, d\ln t_{kl,m}}_{\text{Modal Substitution}} + \underbrace{\epsilon_{t_{kl}}^{\Xi_{kl}}\epsilon_{t_{kl,m}}^{t_{kl}}\, d\ln t_{kl,m}}_{\text{Modal Complementarity}}
$$

Starting with the expression for mode-specific traffic, we obtain,

$$
\Xi_{kl,m} = \frac{t_{kl,m}^{-\eta}}{t_{kl}^{-\eta}}\times\Xi_{kl}
$$

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

Totally differentiating, we obtain,

$$
d\ln\Xi_{kl,m} = \epsilon_{t_{kl,m}^m}^{\pi_{kl}^m}\, d\ln t_{kl,m} + \epsilon_{t_{kl}}^{\Xi_{kl}}\epsilon_{t_{kl,m}}^{t_{kl}}\, d\ln t_{kl,m}
$$

Totally differentiating the mode choice share, π_{kl}^m, we obtain,

$$
\epsilon_{t_{kl,m}}^{\pi_{kl}^m} = -\eta\left(1-\omega_{ik,m}\right)
$$

$$
\epsilon_{t_{kl,m'}}^{\pi_{kl}^m} = \eta\left(\omega_{ik,m'}\right)
$$

Totally differentiating the aggregate traffic equation, we obtain,

$$
\epsilon_{t_{kl}}^{\Xi_{kl}} = \left(-\theta+\epsilon_{t_{kl}}^{P_k^{-\theta}}+\epsilon_{t_{kl}}^{\Pi_l^{-\theta}}\right)
$$

Separating out direct partial equilibrium effects from indirect general equilibrium effects,

$$
\epsilon_{t_{kl}}^{\Xi_{kl},PE} = -\theta
$$

$$
\epsilon_{t_{kl}}^{\Xi_{kl},GE} = \left(\epsilon_{t_{kl}}^{P_k^{-\theta}}+\epsilon_{t_{kl}}^{\Pi_l^{-\theta}}\right)
$$

And finally, totally differentiating the expression for aggregate transport costs, we obtain,

$$d\ln t_{ik} \propto -\frac{1}{\eta} d\ln\left(\sum_{m\in\mathscr{M}(i,k)} t_{ik,m}^{-\eta}\right)$$
$$= \sum_{m\in\mathscr{M}(i,k)} \frac{t_{ik,m}^{-\eta}}{\sum_{m\in\mathscr{M}(i,k)} t_{ik,m}^{-\eta}} d\ln\left(t_{ik,m}\right)$$

which implies,

$$\epsilon_{t_{kl,m}}^{t_{kl}} = \omega_{ik,m}$$

Solving for the market access elasticities, from the recursive equilibrium equations,

$$\Pi_i^{-\theta_i} = \left(t_{ii}^{-\theta_i}\right)\frac{\delta_i}{P_i^{-\theta_i}} + \sum_{k\in\mathscr{F}(i)} t_{ik}^{-\theta_i}\tilde{\Pi}_k^{-\theta_i}$$

$$P_j^{-\theta_j} = \left(t_{jj}^{-\theta_j}\right)\frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{k\in\mathscr{B}(j)} t_{kj}^{-\theta_j} P_k^{-\theta_j}$$

For the symmetric case, $\theta_i = \theta$ and totally differentiating, we obtain,

$$\frac{dP_j^{-\theta}}{P_j^{-\theta}} = \left(t_{jj}^{-\theta}\right)P_j^{\theta}\frac{\gamma_j}{\Pi_j^{-\theta_j}}\cdot\frac{d\gamma_j}{\gamma_j} + \left(t_{jj}^{-\theta}\right)P_j^{\theta}\frac{\gamma_j}{\Pi_j^{-\theta_j}}\frac{d\Pi_i^{-\theta}}{\Pi_i^{-\theta}} + \sum_{k\in\mathscr{B}(j)}\frac{t_{kj}^{-\theta}P_k^{-\theta}}{P_j^{-\theta}}\cdot\frac{dP_k^{-\theta}}{P_k^{-\theta}} - \theta\sum_{k\in\mathscr{B}(j)}\frac{t_{kj}^{-\theta}P_k^{-\theta}}{P_j^{-\theta}}\frac{dt_{kj}}{t_{kj}}$$

$$\frac{d\Pi_i^{-\theta}}{\Pi_i^{-\theta}} = \left(t_{ii}^{-\theta}\right)\Pi_i^{\theta}\frac{\delta_i}{P_i^{-\theta_i}}\frac{d\delta_i}{\delta_i} + \left(t_{ii}^{-\theta}\right)\Pi_i^{\theta}\frac{\delta_i}{P_i^{-\theta_i}}\frac{dP_j^{-\theta}}{P_j^{-\theta}} + \sum_{k\in\mathscr{F}(i)}\frac{t_{ik}^{-\theta}\Pi_k^{-\theta}}{\Pi_i^{-\theta}}\frac{d\Pi_k^{-\theta}}{\Pi_k^{-\theta}} - \theta\sum_{k\in\mathscr{F}(i)}\frac{t_{ik}^{-\theta}\Pi_k^{-\theta}}{\Pi_i^{-\theta}}\frac{dt_{ki}}{t_{ki}}$$

In logs,

$$d\ln P_j^{-\theta} = \left(t_{jj}^{-\theta}\right)P_j^{\theta}\frac{\gamma_j}{\Pi_j^{-\theta_j}}\cdot d\ln\gamma_j + \left(t_{jj}^{-\theta}\right)P_j^{\theta}\frac{\gamma_j}{\Pi_j^{-\theta_j}}d\ln\Pi_i^{-\theta} + \sum_{k\in\mathscr{B}(j)}\frac{t_{kj}^{-\theta}P_k^{-\theta}}{P_j^{-\theta}}\cdot d\ln P_k^{-\theta} - \theta\sum_{k\in\mathscr{B}(j)}\frac{t_{kj}^{-\theta}P_k^{-\theta}}{P_j^{-\theta}}d\ln t_{kj}$$

$$d\ln\Pi_i^{-\theta} = \left(t_{ii}^{-\theta}\right)\Pi_i^{\theta}\frac{\delta_i}{P_i^{-\theta_i}}d\ln\delta_i + \left(t_{ii}^{-\theta}\right)\Pi_i^{\theta}\frac{\delta_i}{P_i^{-\theta_i}}d\ln P_j^{-\theta} + \sum_{k\in\mathscr{F}(i)}\frac{t_{ik}^{-\theta}\Pi_k^{-\theta}}{\Pi_i^{-\theta}}d\ln\Pi_k^{-\theta} - \theta\sum_{k\in\mathscr{F}(i)}\frac{t_{ik}^{-\theta}\Pi_k^{-\theta}}{\Pi_i^{-\theta}}d\ln t_{ki}$$

In matrix notation, we have,

$$dP = \Omega^{i,P}d\gamma + \Omega^{i,P}d\Pi + \Omega^{t,P}dP - \theta\Omega^{t,P}dt$$
$$d\Pi = \Omega^{i,\Pi}d\gamma + \Omega^{i,\Pi}dP + \Omega^{t,\Pi}d\Pi - \theta\Omega^{t,\Pi}dt$$

where $dP = \left[d\ln P_j^{-\theta}\right]$ and $d\Pi = \left[d\ln\Pi_j^{-\theta}\right]$ are the vectors of log changes in market access indices, while $\Omega^{i,\Pi} = \left[\left(t_{ii}^{-\theta}\right)\Pi_i^{\theta}\frac{\delta_i}{P_i^{-\theta_i}}\right]_{ii}$ and $\Omega^{i,P} = \left[\left(t_{jj}^{-\theta}\right)P_j^{\theta}\frac{\gamma_j}{\Pi_j^{-\theta_j}}\right]_{jj}$ are diagonal matrices and $\Omega^{t,P} = \left[\frac{t_{kj}^{-\theta}P_k^{-\theta}}{P_j^{-\theta}}\right]_{kj}$ and $\Omega^{t,\Pi} = \left[\frac{t_{ik}^{-\theta}\Pi_k^{-\theta}}{\Pi_i^{-\theta}}\right]_{ik}$ are full matrices that determine the propagation of tran sport cost changes to changes in price indices. Solving for equilibrium adjustments, we obtain,

$$d\Pi = \left(\mathbf{I} - \Omega^{t,\Pi}\right)^{-1}\Omega^{i,\Pi}d\gamma + \left(\mathbf{I} - \Omega^{t,\Pi}\right)^{-1}\Omega^{i,\Pi}dP - \theta\left(\mathbf{I} - \Omega^{t,\Pi}\right)^{-1}\Omega^{t,\Pi}dt$$
$$dP = \left(\mathbf{I} - \Omega^{t,P}\right)^{-1}\Omega^{i,P}d\gamma + \left(\mathbf{I} - \Omega^{t,P}\right)^{-1}\Omega^{i,P}d\Pi - \theta\left(\mathbf{I} - \Omega^{t,P}\right)^{-1}\Omega^{t,P}dt$$

Define the centrality terms,

$$\tilde{\Omega}^{t,\Pi} \equiv \left(\mathbf{I} - \Omega^{t,\Pi}\right)^{-1}\Omega^{t,\Pi}$$

$$\tilde{\Omega}^{t,P} \equiv \left(\mathbf{I} - \Omega^{t,P}\right)^{-1}\Omega^{t,P}$$

we obtain,

$$d\Pi = \left(\mathbf{I} - \Omega^{t,\Pi}\right)^{-1} \Omega^{i,\Pi} d\gamma + \left(\mathbf{I} - \Omega^{t,\Pi}\right)^{-1} \Omega^{i,\Pi} dP - \theta \tilde{\Omega}^{t,\Pi} dt$$

$$dP = \left(\mathbf{I} - \Omega^{t,P}\right)^{-1} \Omega^{i,P} d\gamma + \left(\mathbf{I} - \Omega^{t,P}\right)^{-1} \Omega^{i,P} d\Pi - \theta \tilde{\Omega}^{t,P} dt$$

In terms of (approximate) elasticities with regard to mode-segment specific transport cost changes, we obtain,

$$\epsilon_{t_{kl}}^{P_k^{-\theta}} \approx -\theta \left[\tilde{\Omega}^{t,P}\right]_{kl}$$

$$\epsilon_{t_{kl}}^{\Pi_l^{-\theta}} \approx -\theta \left[\tilde{\Omega}^{t,\Pi}\right]_{kl}$$

which illustrates that the impact of a transport cost change depends on the centrality of segment kl to the overall transportation network.

We now turn towards the relative magnitude of modal substitution and complementarity. We have,

$$\begin{aligned}
\frac{d \ln \Xi_{kl,m}}{d \ln t_{kl,m'}} &= \epsilon_{t_{kl,m'}^m}^{\pi_{kl}^m} + \epsilon_{t_{kl}}^{\Xi_{kl}} \epsilon_{t_{kl,m}}^{t_{kl}} \\
&= \epsilon_{t_{kl,m'}^m}^{\pi_{kl}^m} + \left(-\theta + \epsilon_{t_{kl}}^{P_k^{-\theta}} + \epsilon_{t_{kl}}^{\Pi_l^{-\theta}}\right) \epsilon_{t_{kl,m'}}^{t_{kl}} \\
&\approx \epsilon_{t_{kl,m'}^m}^{\pi_{kl}^m} - \theta \left(1 + \left[\tilde{\Omega}^{t,P}\right]_{kl} + \left[\tilde{\Omega}^{t,\Pi}\right]_{kl}\right) \epsilon_{t_{kl,m'}}^{t_{kl}} \\
&= \eta \left(\omega_{kl,m'}\right) - \theta \left(1 + \left[\tilde{\Omega}^{t,P}\right]_{kl} + \left[\tilde{\Omega}^{t,\Pi}\right]_{kl}\right) \omega_{ik,m'}
\end{aligned}$$

We can examine the relative absolute magnitude of the two different channels,

$$\begin{aligned}
\eta + \theta &\lessgtr \theta \left(\left[\tilde{\Omega}^{t,P}\right]_{kl} + \left[\tilde{\Omega}^{t,\Pi}\right]_{kl}\right) \\
\left(1 + \frac{\eta}{\theta}\right) &\lessgtr \left(\tilde{\Omega}_{kl}^{t,P} + \tilde{\Omega}_{kl}^{t,\Pi}\right)
\end{aligned} \tag{25}$$

which formalizes the insight that if the general equilibrium elasticities exceed the unit elasticity plus the relative magnitude between the partial equilibrium substitution and aggregate trade effect, then the complementarity effect dominates. This is the case if the centrality of the transport segment is particularly high. In conclusion, we have the following, which gives us Equation (16),

$$\frac{d \ln \Xi_{kl,m}}{d \ln t_{kl,m'}} \approx \theta \omega_{ik,m'} \left[\frac{\eta}{\theta} - 1 + \left[\tilde{\Omega}^{t,P}\right]_{kl} + \left[\tilde{\Omega}^{t,\Pi}\right]_{kl}\right] = \left[\underbrace{\eta}_{\substack{\text{Direct Substitution} \\ \text{Effect}}} \quad \underbrace{-\theta \left(1 + \left[\tilde{\Omega}^{t,P}\right]_{kl} + \left[\tilde{\Omega}^{t,\Pi}\right]_{kl}\right)}_{\substack{\text{Indirect Complementarity} \\ \text{Effect}}}\right] \omega_{kl,m'}$$

A.8 Derivations for Section 5.1: Proofs of Proposition 1

This section presents the proof for Proposition 1. We proceed in two steps. In a first step we derive the change in the equilibrium conditions in terms of market access terms before then substitution the model specific elements.

A.8.1 Preliminaries

We can write equilibrium trade flows as,

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

where γ_i and δ_j are cumulative flows out of and origin and into a destination, respectively, and Π_i and P_j are origin and destination market access terms. Given the recursive routing formulation, trade costs can be represented as:

$$\tau_{ij} = \left(\sum_{k \in \mathcal{F}(i)} \left(t_{ik} \tau_{kj} \right)^{-\theta} \right)^{-\frac{1}{\theta}}$$

And furthermore from the nested choice along the multi-layered graph, we have,

$$t_{ik}^{-\theta} = \left(\sum_{m \in \mathcal{M}(i,k)} \tilde{t}_{ik,m}^{-\eta} \right)^{\frac{\theta}{\eta}|}$$

In terms of market clearing and trade balance we have the following equilibrium conditions,

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

A.8.2 Deriving the equilibrium equation

Starting with the first equilibrium condition we have:

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

$$\gamma_i = \sum_j \tau_{ij}^{-\theta_j} \times \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}} \times \frac{\delta_j}{P_j^{-\theta_j}}$$

$$\Pi_i^{-\theta_i} = \sum_j \tau_{ij}^{-\theta_j} \times \Pi_i^{\theta_j \frac{\theta_j}{\theta_i} - \theta_i} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \times \frac{\delta_j}{P_j^{-\theta_j}}$$

$$\Pi_i^{-\theta_i} = \sum_j \tau_{ij}^{-\theta_j} \times \Pi_i^{\Delta_{i,j}} \times (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \times \frac{\delta_j}{P_j^{-\theta_j}}$$

$$\Pi_i^{-\theta_i} = \tau_{ii}^{-\theta_i} \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{j \neq i} \tau_{ij}^{-\theta_j} \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \frac{\delta_j}{P_j^{-\theta_j}}$$

$$\Pi_i^{-\theta_i} = \tau_{ii}^{-\theta_i} \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{j \neq i} \left(\sum_{k \in \mathscr{B}(k)} t_{ik}^{-\theta_j} \tau_{kj}^{-\theta_j} \right) \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \frac{\delta_j}{P_j^{-\theta_j}}$$

$$\Pi_i^{-\theta_i} = \tau_{ii}^{-\theta_i} \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{j \neq i} \left(\sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta_j} \tau_{kj}^{-\theta_j} \right) \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \frac{\delta_j}{P_j^{-\theta_j}}$$

$$\Pi_i^{-\theta_i} = \tau_{ii}^{-\theta_i} \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta_i} \left(-\tau_{ki}^{-\theta_i} \frac{\delta_i}{p_i^{-\theta_i}} + \sum_j t_{ik}^{\Delta\theta_j} \tau_{kj}^{-\theta_j} \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \frac{\delta_j}{P_j^{-\theta_j}} \right)$$

$$\Pi_i^{-\theta_i} = \left(t_{ii}^{-\theta_i} + \sum_{k \in \mathscr{F}(i)} (t_{ik} \tau_{ki})^{-\theta_i} \right) \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta_i} \left(-\tau_{ki}^{-\theta_i} \frac{\delta_i}{p_i^{-\theta_i}} + \sum_j t_{ik}^{\Delta\theta_j} \tau_{kj}^{-\theta_j} \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \frac{\delta_j}{P_j^{-\theta_j}} \right)$$

$$\Pi_i^{-\theta_i} = \left(t_{ii}^{-\theta_i} \right) \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta_i} \left(\sum_j t_{ik}^{\Delta\theta_j} \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \tau_{kj}^{-\theta_j} \frac{\delta_j}{P_j^{-\theta_j}} \right)$$

$$\Pi_i^{-\theta_i} = t_{ii}^{-\theta_i} \frac{\delta_i}{P_i^{-\theta_i}} + \sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta_i} \tilde{\Pi}_k^{-\theta_i}$$

where in the second line we defined, $\Pi_i^{\Delta_{i,j}} \equiv \Pi_i^{\theta_j \frac{\theta_j}{\theta_i} - \theta_i}$ and where $\tilde{\Pi}_k^{-\theta_i} = \left(\sum_j t_{ik}^{\Delta\theta_j} \Pi_i^{\Delta_{i,j}} (\gamma_i)^{\frac{\theta_j}{\theta_i} - 1} \tau_{kj}^{-\theta_j} \frac{\delta_j}{P_j^{-\theta_j}} \right)^{\frac{\theta_i}{\theta_j}}$ is the approximate price index that controls for differences in elasticities of substitution between neighboring nodes. And where in the fourth line we have used the fact that at the penultimate node the set of predecessor nodes of k ($\mathscr{B}(k)$) coincides with set of successor nodes of i ($\mathscr{F}(i)$).

Continuing with the second equilibrium condition,

$$\delta_j = \sum_i X_{ij}$$

$$\delta_j = \sum_i \tau_{ij}^{-\theta_j} \times \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}} \times \frac{\delta_j}{P_j^{-\theta_j}}$$

$$P_j^{-\theta_j} = \sum_i \tau_{ij}^{-\theta_j} \times \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}}$$

$$P_j^{-\theta_j} = \tau_{jj}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{i \neq j} \tau_{ij}^{-\theta_j} \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}}$$

$$P_j^{-\theta_j} = \tau_{jj}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{i \neq j} \left(\sum_{k \in \mathscr{B}(j)} \tau_{ik}^{-\theta_j} t_{kj}^{-\theta_j} \right) \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}}$$

$$P_j^{-\theta_j} = \tau_{jj}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta}} + \sum_{k \in \mathscr{B}(j)} t_{kj}^{-\theta_j} \left(-\tau_{jk}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_i \tau_{ik}^{-\theta_j} \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}} \right)$$

$$P_j^{-\theta_j} = \left(t_{jj}^{-\theta_j} + \sum_{k \in \mathscr{B}(j)} \left(t_{jk} \tau_{kj} \right)^{-\theta_j} \right) \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{k \in \mathscr{B}(j)} t_{kj}^{-\theta_j} \left(-\tau_{jk}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_i \tau_{ik}^{-\theta_j} \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}} \right)$$

$$P_j^{-\theta_j} = \left(t_{jj}^{-\theta_j} \right) \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{k \in \mathscr{B}(j)} t_{kj}^{-\theta_j} \left(\sum_i \tau_{ik}^{-\theta_j} \left(\frac{\gamma_i}{\Pi_i^{-\theta_j}} \right)^{\frac{\theta_j}{\theta_i}} \right)$$

$$P_j^{-\theta_j} = t_{jj}^{-\theta_j} \frac{\gamma_j}{\Pi_j^{-\theta_j}} + \sum_{k \in \mathscr{B}(j)} t_{kj}^{-\theta_j} P_k^{-\theta_j}$$

where in the last line we used the definition of the recursive transport cost (Equation (1)), $\tau_{ii}^{-\theta} = \left(t_{ii}^{-\theta} + \sum_{k \in \mathscr{F}(i)} \left(t_{ik} \tau_{kj} \right)^{-\theta} \right)$. In what follows, we derive the counterfactual equation for the special case with symmetric substitution elasticities, ($\theta_i = \theta \quad \forall i$), i.e.

$$\Pi_i^{-\theta} = \left(t_{ii}^{-\theta} \right) \frac{\delta_i}{P_i^{-\theta}} + \sum_{k \in \mathscr{F}(i)} t_{ik}^{-\theta} \Pi_k^{-\theta}$$

$$P_j^{-\theta} = t_{jj}^{-\theta} \frac{\gamma_j}{\Pi_j^{-\theta}} + \sum_{k \in \mathscr{B}(j)} t_{kj}^{-\theta} P_k^{-\theta}$$

We can make the relationship between the mode-specific transportation cost and the spatial aggregate equilibrium more explicit by substitution the definition of the edge-specific transport cost in terms of the multi-layered mode specific transport cost, i.e.

$$P_i^{-\theta} = t_{ii}^{-\theta} \frac{\gamma_i}{\Pi_i^{-\theta}} + \sum_{k \in \mathscr{F}(j)} \left(\sum_{m \in \mathscr{M}(k,i)} \tilde{t}_{ki,m}^{-\eta} \right)^{\frac{\theta}{\eta}|} P_k^{-\theta}$$

$$= t_{ii}^{-\theta} \frac{\gamma_i}{\Pi_i^{-\theta}} + \sum_{k \in \mathscr{F}(j)} \left(\sum_{m \in \mathscr{M}(k,i)} \tilde{t}_{ki,m}^{-\eta} P_k^{-\eta} \right)^{\frac{\theta}{\eta}|}$$

$$\Pi_i^{-\theta} = t_{ii}^{-\theta}\frac{\delta_i}{P_i^{-\theta}} + \sum_{k\in\mathscr{F}(i)}\left(\sum_{m\in\mathscr{M}(i,k)}\tilde{t}_{ik,m}^{-\eta}\right)^{\frac{\theta}{\eta}|}\Pi_k^{-\theta}$$
$$= t_{ii}^{-\theta}\frac{\delta_i}{P_i^{-\theta}} + \sum_{k\in\mathscr{F}(i)}\left(\sum_{m\in\mathscr{M}(i,k)}\tilde{t}_{ik,m}^{-\eta}\Pi_k^{-\eta}\right)^{\frac{\theta}{\eta}|}$$

where we notice that the equilibrium equations then inherit a the structure of nested choice models directly and where the modal substitution parameter than governs the impact of mode-specific changes on aggregate endogenous variables.

A.8.3 Deriving the equilibrium system in changes

Start with (recursive) equilibrium condition from the previous subsection,

$$\Pi_i^{-\theta} = t_{ii}^{-\theta}\frac{\delta_i}{P_i^{-\theta}} + \sum_{k\in\mathscr{F}(i)} t_{ik}^{-\theta}\Pi_k^{-\theta}$$

$$P_i^{-\theta} = t_{ii}^{-\theta}\frac{\gamma_i}{\Pi_i^{-\theta}} + \sum_{k\in\mathscr{F}(j)} t_{ki}^{-\theta}P_k^{-\theta}$$

Expressed in changes,

$$\hat{\Pi}_i^{-\theta} = \left(\frac{t_{ii}^{-\theta}\frac{\delta_i}{P_i^{-\theta}}}{t_{ii}^{-\theta}\frac{\delta_i}{P_i^{-\theta}} + \sum_{k\in\mathscr{F}(i)} t_{ik}^{-\theta}\Pi_k^{-\theta}}\right)\frac{\hat{\delta}_i}{\hat{P}_i^{-\theta}} + \sum_{k\in\mathscr{F}(i)}\left(\frac{t_{ik}^{-\theta}\Pi_k^{-\theta}}{t_{ii}^{-\theta}\frac{\delta_i}{P_i^{-\theta}} + \sum_{k\in\mathscr{F}(i)} t_{ik}^{-\theta}\Pi_k^{-\theta}}\right)\hat{t}_{ik}^{-\theta}\hat{\Pi}_k^{-\theta}$$

and,

$$\hat{P}_i^{-\theta} = \left(\frac{t_{ii}^{-\theta}\frac{\gamma_i}{\Pi_i^{-\theta}}}{t_{ii}^{-\theta}\frac{\gamma_i}{\Pi_i^{-\theta}} + \sum_{k\in\mathscr{F}(j)} t_{ki}^{-\theta}P_k^{-\theta}}\right)\frac{\hat{\gamma}_i}{\hat{\Pi}_i^{-\theta}} + \sum_{k\in\mathscr{F}(i)}\left(\frac{t_{ki}^{-\theta}P_k^{-\theta}}{t_{ii}^{-\theta}\frac{\gamma_i}{\Pi_i^{-\theta}} + \sum_{k\in\mathscr{F}(i)} t_{ki}^{-\theta}P_k^{-\theta}}\right)\hat{t}_{ki}^{-\theta}\hat{P}_k^{-\theta}$$

Multiplying both numerator and denominator by their appropriate market access terms, we obtain,

$$\hat{\Pi}_i^{-\theta} = \left(\frac{\delta_i}{\delta_i + \sum_{k\in\mathscr{F}(i)} t_{ik}^{-\theta}P_k^{-\theta}\Pi_j^{-\theta}}\right)\frac{\hat{\delta}_i}{\hat{P}_i^{-\theta}} + \sum_{k\in\mathscr{F}(i)}\left(\frac{t_{ik}^{-\theta}\Pi_j^{-\theta}P_i^{-\theta}}{\delta_i + \sum_{k\in\mathscr{F}(i)} t_{ik}^{-\theta}P_k^{-\theta}\Pi_j^{-\theta}}\right)\hat{t}_{ik}^{-\theta}\hat{\Pi}_k^{-\theta}$$

$$\hat{P}_i^{-\theta} = \left(\frac{\gamma_i}{\gamma_i + \sum_{k\in\mathscr{F}(i)} t_{ki}^{-\theta}P_k^{-\theta}\Pi_i^{-\theta}}\right)\frac{\hat{\gamma}_i}{\hat{\Pi}_i^{-\theta}} + \sum_{k\in\mathscr{F}(i)}\left(\frac{t_{ki}^{-\theta}P_k^{-\theta}\Pi_i^{-\theta}}{\gamma_i + \sum_{k\in\mathscr{F}(i)} t_{ki}^{-\theta}P_k^{-\theta}\Pi_i^{-\theta}}\right)\hat{t}_{ki}^{-\theta}\hat{P}_k^{-\theta}$$

Simplifying we obtain,

$$\hat{\Pi}_i^{-\theta} = \left(\frac{\delta_i}{\delta_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\frac{\hat{\delta}_i}{\hat{P}_i^{-\theta}} + \sum_{k\in\mathscr{F}(i)}\left(\frac{\Xi_{ik}}{\delta_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{ik}^{-\theta}\hat{\Pi}_k^{-\theta}$$

$$\hat{P}_i^{-\theta} = \left(\frac{\gamma_i}{\gamma_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\frac{\hat{\gamma}_i}{\hat{\Pi}_i^{-\theta}} + \sum_{k\in\mathscr{F}(i)}\left(\frac{\Xi_{ki}}{\gamma_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{ki}^{-\theta}\hat{P}_k^{-\theta}$$

Next, we proceed by deriving the nested transport equilibrium at the edge-level. Recall, that given the nested choice above the changes in aggregate transport cost is given by,

$$\hat{t}_{ik}^{-\theta} = \left(\sum_{m\in\mathscr{M}(i,k)}\frac{\left(t'_{ik,m}\right)^{-\eta}}{t_{ik}^{-\eta}}\right)^{\frac{\theta}{\eta}}$$

Multiplying with the appropriate market access terms, we obtain,

$$\hat{t}_{ik}^{-\theta} = \left(\sum_{m\in\mathscr{M}(i,k)}\frac{\Xi_{ik.m}}{\Xi_{ik}}\hat{t}_{ik,m}^{-\eta}\right)^{\frac{\theta}{\eta}}$$

For the changes in transport cost we have,

$$\hat{t}_{kl,1} = \hat{\hat{t}}_{kl,0}^{\frac{1}{1+\eta\lambda_0}} \times \hat{t}_{kl}^{\frac{\lambda_0(\eta-\theta)}{1+\eta\lambda_0}} \times \hat{P}_k^{\frac{-\theta\lambda_0}{1+\eta\lambda_0}} \times \hat{\Pi}_l^{\frac{-\theta\lambda_0}{1+\eta\lambda_0}}$$

And second, we have for any non-primary transport mode,

$$\hat{t}_{kl,m} = \hat{\hat{s}}_{kk,m}\hat{\tau}_{kl,m}\hat{\hat{s}}_{ll,m}\hat{P}_k^{-\theta\lambda_m}\hat{\Pi}_{k,m}^{-\theta\lambda_m}\hat{\Pi}_l^{-\theta\lambda_m}\hat{P}_{l,m}^{-\theta\lambda_m}$$

For the trade model here we have,

$$\delta_i = E_i$$
$$\gamma_i = Y_i$$

And for the price indices in changes we have,

$$\hat{P}_i = \hat{y}_i\hat{l}_i^{\beta-1}\hat{W}^{-1}$$
$$\hat{\Pi}_i = \hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}$$

and furthermore, we have the mode specific market access terms,

$$\hat{P}_{l,m}^{-\theta} \equiv \sum_{k\in\mathscr{B}(l)}\frac{\Xi_{kl,m}}{\sum_{k\in\mathscr{B}(l)}\Xi_{kl,m}}\hat{t}_{kl,m}^{-\eta} \times \hat{t}_{kl}^{\eta-\theta} \times \hat{P}_k^{-\theta}$$

$$\hat{\Pi}_{k,m}^{-\theta} \equiv \sum_{l\in\mathscr{F}(k)}\frac{\Xi_{kl,m}}{\sum_{l\in\mathscr{F}(k)}\Xi_{kl,m}} \times \hat{t}_{kl,m}^{-\eta} \times \hat{t}_{kl}^{\eta-\theta} \times \hat{\Pi}_l^{-\theta}$$

Given an initial guess for the market access terms, we can solve for the aggregate transport costs, $\hat{t}_{ik}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ik,1},\hat{t}_{ik,m}\right)$, by iteratively updating the following equation until convergence,

$$\begin{aligned}\hat{t}_{ik}^{-\theta} &= \left(\frac{\Xi_{ik.0}}{\Xi_{ik}}\hat{t}_{ik,0}^{-\eta} + \sum_{m\neq 0}\frac{\Xi_{ik.m}}{\Xi_{ik}}\hat{t}_{ik,m}^{\eta}\right)^{\frac{\theta}{\eta}} \\ &= \left(\frac{\Xi_{ik.0}}{\Xi_{ik}}\left[\hat{\hat{t}}_{ik,0}^{\frac{1}{+\eta\eta\lambda_0}} \times \hat{t}_{ik}^{\frac{\lambda_0(\eta-\theta)}{1+\eta\lambda_0}} \times \hat{P}_i^{\frac{-\theta\lambda_0}{1+\eta\lambda_0}} \times \hat{\Pi}_k^{\frac{-\theta\lambda_0}{1+\eta\lambda_0}}\right]^{-\eta} + \sum_{m\neq 0}\frac{\Xi_{ik.m}}{\Xi_{ik}}\left[\hat{\hat{s}}_{ii,m}\hat{\tau}_{ik,m}\hat{\hat{s}}_{kk,m}\hat{P}_i^{-\theta\lambda_m}\hat{\Pi}_{i,m}^{-\theta\lambda_m}\hat{\Pi}_k^{-\theta\lambda_m}\hat{P}_{k,m}^{-\theta\lambda_m}\right]^{-\eta}\right)^{\frac{\theta}{\eta}}\end{aligned}$$

Writing out and substituting for the transport cost terms,

$$\hat{P}_i^{-\theta}\hat{\Pi}_i^{-\theta} = \left(\frac{\delta_i}{\delta_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{\delta}_i + \sum_{k\in\mathscr{F}(i)}\left(\frac{\Xi_{ik}}{\delta_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{ik}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{\tilde{t}}_{ik,0},\hat{\tilde{t}}_{ik,m}\right)\hat{P}_i^{-\theta}\hat{\Pi}_k^{-\theta}$$

$$\hat{P}_i^{-\theta}\hat{\Pi}_i^{-\theta} = \left(\frac{\gamma_i}{\gamma_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{\gamma}_i + \sum_{k\in\mathscr{F}(i)}\left(\frac{\Xi_{ki}}{\gamma_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{ki}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{\tilde{t}}_{ki,0},\hat{\tilde{t}}_{ki,m}\right)\hat{P}_k^{-\theta}\hat{\Pi}_i^{-\theta}$$

Substituting the price indices in changes, we get,

$$\begin{aligned}\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_i\hat{l}_i^{\beta-1}\hat{W}^{-1}\right)^{-\theta} &= \left(\frac{E_i}{E_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{y}_i \\ &+ \sum_{k\in\mathscr{F}(i)}\left(\frac{\Xi_{ik}}{E_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{ik}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ik,0},\hat{t}_{ik,m}\right)\left(\hat{y}_i\hat{l}_i^{\beta-1}\hat{W}^{-1}\right)^{-\theta}\left(\hat{l}_k^{\alpha+1}\hat{y}_k^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\end{aligned}$$

$$\begin{aligned}\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_i\hat{l}_i^{\beta-1}\hat{W}^{-1}\right)^{-\theta} &= \left(\frac{Y_i}{Y_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{y}_i \\ &+ \sum_{k\in\mathscr{F}(i)}\left(\frac{\Xi_{ki}}{Y_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{ki}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ki,0},\hat{t}_{ki,m}\right)\left(\hat{y}_k\hat{l}_k^{\beta-1}\hat{W}^{-1}\right)^{-\theta}\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\end{aligned}$$

Simplifying and defining, $\hat{\chi} = \hat{W}^{-\theta}$, we have,

$$\hat{y}_i^{1+\theta}\hat{l}_i^{-\theta(1+\alpha)} = \hat{\chi}\left(\frac{E_i}{E_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{y}_i^{1+\theta}\hat{l}_i^{-\theta(1-\beta)} + \sum_{k\in\mathscr{F}(i)}\left(\frac{\Xi_{ik}}{E_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{ik}^{-\theta}\hat{y}_k^{1+\theta}\hat{l}_k^{-\theta(1+\alpha)}$$

$$\hat{y}_i^{-\theta}\hat{l}_i^{\theta(1-\beta)} = \hat{\chi}\left(\frac{Y_i}{Y_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{y}_i^{-\theta}\hat{l}_i^{\theta(1+\alpha)} + \sum_{k\in\mathscr{F}(i)}\left(\frac{\Xi_{ki}}{Y_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{ki}^{-\theta}\hat{y}_k^{-\theta}\hat{l}_k^{\theta(1-\beta)}$$

A.9 Derivations for Section 5.3: Predicted versus Observed Transport Mode-level Trade Flows

For aggregate traffic we have,

$$\boldsymbol{\Xi} = \mathbf{PA}\Pi$$

Inverting,

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

For the traffic equations we have,

$$\boldsymbol{\Xi}_m = \mathbf{PT_m}\Pi$$

Inverting,

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

Bilateral mode specific shares are given by,

$$\begin{aligned}\pi_{ij,k,m} &= \frac{(w_i/A_i)^{-\theta}(t_{ik}\tau_{ki})^{-\theta}}{\sum_{i\in\mathscr{N}}(w_i/A_i)^{-\theta}\sum_{k'\in\mathscr{F}(i)}(t_{ik'}\tau_{k'i})^{-\theta}}\frac{\tilde{t}_{ik,m}^{-\eta}}{t_{ik}} \\ &= \pi_{ij,k}\times\pi_{ij,k}^{m}\end{aligned}$$

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

$$X_{ij} = \left(\sum_{k \in \mathcal{N}} t_{ik}^{-\theta} \tau_{kj}^{-\theta} \right) \times \frac{\gamma_i}{\Pi_i^{-\theta}} \times \frac{\delta_j}{P_j^{-\theta}}$$
$$= \sum_{k \in \mathcal{N}} \underbrace{t_{ik}^{-\theta} \times \tau_{kj}^{-\theta} \times \frac{\gamma_i}{\Pi_i^{-\theta}} \times \frac{\delta_j}{P_j^{-\theta}}}_{\equiv X_{ij,k}}$$

$$X_{ij} = \left(\sum_{k \in \mathcal{N}} \sum_{l \in \mathcal{N}(k)} \tau_{ik}^{-\theta} t_{kl}^{-\theta} \tau_{lj}^{-\theta} \right) \times \frac{\gamma_i}{\Pi_i^{-\theta}} \times \frac{\delta_j}{P_j^{-\theta}}$$
$$= \sum_{k \in \mathcal{N}} \sum_{l \in \mathcal{N}(k)} \underbrace{\tau_{ik}^{-\theta} t_{kl}^{-\theta} \tau_{lj}^{-\theta} \times \frac{\gamma_i}{\Pi_i^{-\theta}} \times \frac{\delta_j}{P_j^{-\theta}}}_{\equiv X_{ij,kl}}$$

The share of flows between ij, along link kl and along mode m,

$$X_{ij,kl} = \left(\tau_{ik} t_{kl} \tau_{lj} \right)^{-\theta} \times \frac{\gamma_i}{\Pi_i^{-\theta}} \times \frac{\delta_j}{P_j^{-\theta}}$$

$$X_{ij,kl,m} = \frac{\tilde{t}_{kl,m}^{-\eta}}{t_{kl}^{-\eta}} \times \left(\tau_{ik} t_{kl} \tau_{lj} \right)^{-\theta} \times \frac{\gamma_i}{\Pi_i^{-\theta}} \times \frac{\delta_j}{P_j^{-\theta}}$$

sum over kl to obtain

$$X_{ij,m} = \sum_{k \in \mathbf{N}} \sum_{l \in \mathcal{N}(k)} \frac{\tilde{t}_{kl,m}^{-\eta}}{t_{kl}^{-\eta}} \times \left(\tau_{ik} t_{kl} \tau_{lj} \right)^{-\theta} \times \frac{\gamma_i}{\Pi_i^{-\theta}} \times \frac{\delta_j}{P_j^{-\theta}}$$
$$X_{ij,m} = \frac{\gamma_i}{\Pi_i^{-\theta}} \frac{\delta_j}{P_j^{-\theta}} \sum_{k \in \mathbf{N}} \tau_{ik}^{-\theta} \sum_{l \in \mathcal{N}(k)} \tau_{lj}^{-\theta} \times \frac{\tilde{t}_{kl,m}^{-\eta}}{t_{kl}^{-\eta}} t_{kl}^{-\theta}$$
$$X_{ij,m} = \frac{\gamma_i}{\Pi_i^{-\theta}} \frac{\delta_j}{P_j^{-\theta}} \sum_{k \in \mathbf{N}} \tau_{ik}^{-\theta} \sum_{l \in \mathcal{N}(k)} \tau_{lj}^{-\theta} \times \frac{\tilde{t}_{kl,m}^{-\eta}}{t_{kl}^{\theta-\eta}}$$

In matrix notation, we have,

$$\mathbf{X}_m = \left(\frac{Y}{\Pi} \right) (\mathbf{I} - \mathbf{A})^{-1} \mathbf{T}_m (\mathbf{I} - \mathbf{A})^{-1} \left(\frac{E}{P} \right)$$

Derivations,

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

Finally, we need an expression for $\mathbf{P}\Pi$,

$$\Pi P = Y + \Xi^T 1$$

and,

$$\Pi P = E + \Xi 1$$

We have,

$$\mathbf{X}_m = (Y)\left(\mathrm{diag}\left(\frac{1}{2}(E+Y)+\frac{1}{2}\left(\Xi 1+\Xi^T 1\right)\right)-\Xi\right)^{-1}\boldsymbol{\Xi}_m\left(\mathrm{diag}\left(\frac{1}{2}(E+Y)+\frac{1}{2}\left(\Xi 1+\Xi^T 1\right)\right)-\Xi\right)^{-1}(E)$$

We finally obtain,

$$X_{ij,m} = Y_i \times E_j \times \sum_{(k,l)}\left[\mathbf{D}^X-\Xi\right]^{-1}_{ik}\left[\boldsymbol{\Xi}_m\right]_{kl}\left[\mathbf{D}^X-\Xi\right]^{-1}_{lj}, \tag{26}$$

where $\mathbf{D}^X$ is the diagonal matrix with diagonal elements given by $\left(\frac{1}{2}(E+Y)+\frac{1}{2}\left(\Xi 1+\Xi^T 1\right)\right)_{ii}$

A.10 Derivations for Section 5.1: International Trade

In this subsection, we extend the model to allow for both domestic locations $i, j \in \mathscr{D}$ and foreign locations $l \in \mathscr{F}$. Domestic locations consumer from both domestic and foreign locations, while foreign locations are modeled to only trade with domestic locations. Furthermore, we make a small open economy assumption such that changes in the domestic economy do not affect the labor allocations in the foreign economy. For the domestic economy we have the following conditions,

$$\gamma_i = \sum_{j \in \mathscr{D}} X_{ij} + \sum_{l \in \mathscr{F}} X_{il}$$
$$\delta_i = \sum_{j \in \mathscr{D}} X_{ji} + \sum_{l \in \mathscr{F}} X_{li}$$

Welfare equalization domestically implies,

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

and

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

For the foreign economy we have that market clearing implies,

$$\gamma_l = \sum_j X_{lj} \quad \forall l \in \mathscr{F}$$
$$\delta_l = \sum_j X_{jl} \quad \forall l \in \mathscr{F}$$

For the price indices for foreign locations we obtain,

$$P_l = \frac{w_l u_l}{W_l} \iff$$
$$P_l = Y_l \bar{u}_l L_l^{\beta-1} W_l^{-1} \implies$$
$$\hat{P}_l = \hat{y}_l \hat{W}_l^{-1}$$

and

$$\Pi_l = A_l L_l Y_l^{-\frac{\theta+1}{\theta}} \iff$$
$$\Pi_l = A_l L_l^{\alpha+1} Y_l^{-\frac{\theta+1}{\theta}} \implies$$
$$\hat{\Pi}_l = \hat{y}_l^{-\frac{\theta+1}{\theta}}$$

Considering now that the economy is situated on a graph with both foreign and domestic nodes, i.e. $\mathscr{G} \equiv (\mathscr{N}, \mathscr{L})$, where $\mathscr{N}$ is the union of domestic, $\mathscr{N}_D$, and foreign nodes, $\mathscr{N}_F$, and equivalently we define domestic and foreign successor nodes, $k \in \mathscr{F}_D(i)$, and $l \in \mathscr{F}_F(i)$.

Re-writing the equilibrium system for both foreign and domestic locations, defining domestic and foreign neighboring

notes

$$\hat{P}_i^{-\theta}\hat{\Pi}_i^{-\theta} = \left(\frac{E_i}{E_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{\delta}_i + \sum_{k\in\mathscr{F}_D(i)}\left(\frac{\Xi_{ik}}{E_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{ik}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ik,1},\hat{t}_{ik,m}\right)\hat{\Pi}_k^{-\theta}\hat{P}_i^{-\theta} + \sum_{l\in\mathscr{F}_F(i)}\left(\frac{\Xi_{il}}{E_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{il}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ik,1},\hat{t}_{ik,m}\right)\hat{P}_i^{-\theta}\hat{\Pi}_l^{-\theta}$$

$$\hat{P}_i^{-\theta}\hat{\Pi}_i^{-\theta} = \left(\frac{Y_i}{Y_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{Y}_i + \sum_{k\in\mathscr{F}_D(i)}\left(\frac{\Xi_{ki}}{Y_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{ki}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ki,1},\hat{t}_{ki,m}\right)\hat{P}_k^{-\theta}\hat{\Pi}_i^{-\theta} + \sum_{l\in\mathscr{F}_F(i)}\left(\frac{\Xi_{li}}{Y_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{li}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ki,1},\hat{t}_{ki,m}\right)\hat{P}_l^{-\theta}\hat{\Pi}_i^{-\theta}$$

And similarly for foreign locations, we have,

$$\hat{\Pi}_l^{-\theta}\hat{P}_l^{-\theta} = \left(\frac{E_l}{E_l + \sum_{k\in\mathscr{F}_D(l)}\Xi_{lk}}\right)\hat{y}_l + \sum_{k\in\mathscr{F}_D(l)}\left(\frac{\Xi_{lk}}{E_l + \sum_{k\in\mathscr{F}_D(l)}\Xi_{lk}}\right)\hat{t}_{lk}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{lk,1},\hat{t}_{lk,m}\right)\hat{P}_l^{-\theta}\hat{\Pi}_k^{-\theta}$$

$$\hat{P}_l^{-\theta}\hat{\Pi}_l^{-\theta} = \left(\frac{Y_l}{Y_l + \sum_{k\in\mathscr{F}_D(l)}\Xi_{kl}}\right)\hat{y}_l + \sum_{k\in\mathscr{F}_D(l)}\left(\frac{\Xi_{kl}}{Y_l + \sum_{k\in\mathscr{F}_D(l)}\Xi_{kl}}\right)\hat{t}_{kl}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{kl,1},\hat{t}_{kl,m}\right)\hat{P}_k^{-\theta}\hat{\Pi}_l^{-\theta}$$

Substituting,

$$\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_i\hat{l}_i^{\beta-1}\hat{W}^{-1}\right)^{-\theta} = \left(\frac{E_i}{E_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{y}_i + \sum_{k\in\mathscr{F}_D(i)}\left(\frac{\Xi_{ik}}{E_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{ik}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ik,1},\hat{t}_{ik,m}\right)\left(\hat{y}_i\hat{l}_i^{\beta-1}\hat{W}^{-1}\right)^{-\theta}\left(\hat{l}_k^{\alpha+1}\hat{y}_k^{-\frac{\theta+1}{\theta}}\right)^{-\theta} + \sum_{l\in\mathscr{F}_F(i)}\left(\frac{\Xi_{il}}{E_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{il}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ik,1},\hat{t}_{ik,m}\right)\left(\hat{y}_i\hat{l}_i^{\beta-1}\hat{W}^{-1}\right)^{-\theta}\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

$$\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_i\hat{l}_i^{\beta-1}\hat{W}^{-1}\right)^{-\theta} = \left(\frac{Y_i}{Y_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{y}_i + \sum_{k\in\mathscr{F}_D(i)}\left(\frac{\Xi_{ki}}{Y_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{ki}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ki,1},\hat{t}_{ki,m}\right)\left(\hat{y}_k\hat{l}_k^{\beta-1}\hat{W}^{-1}\right)^{-\theta}\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta} + \sum_{l\in\mathscr{F}_F(i)}\left(\frac{\Xi_{li}}{Y_i + \sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{li}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ki,1},\hat{t}_{ki,m}\right)\left(\hat{y}_l\hat{W}_l^{-1}\right)^{-\theta}\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

$$\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_l\hat{W}_l^{-1}\right)^{-\theta}=\left(\frac{E_l}{E_l+\sum_{k\in\mathscr{F}_D(l)}\Xi_{lk}}\right)\hat{y}_l+\sum_{k\in\mathscr{F}_D(l)}\left(\frac{\Xi_{lk}}{E_l+\sum_{k\in\mathscr{F}_D(l)}\Xi_{lk}}\right)\hat{t}_{lk}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{lk,1},\hat{t}_{lk,m}\right)\left(\hat{y}_l\hat{W}_l^{-1}\right)^{-\theta}\left(\hat{l}_k^{\alpha+1}\hat{y}_k^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

$$\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_l\hat{W}_l^{-1}\right)^{-\theta}=\left(\frac{Y_l}{Y_l+\sum_{k\in\mathscr{F}_D(l)}\Xi_{kl}}\right)\hat{y}_l+\sum_{k\in\mathscr{F}_D(l)}\left(\frac{\Xi_{kl}}{Y_l+\sum_{k\in\mathscr{F}_D(l)}\Xi_{kl}}\right)\hat{t}_{kl}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{kl,1},\hat{t}_{kl,m}\right)\left(\hat{y}_k\hat{l}_k^{\beta-1}\hat{W}^{-1}\right)^{-\theta}\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

Multiplying both sides by $\hat{W}^{-\theta}$, and applying the definition $\hat{\chi}=\hat{W}^{-\theta}$,

$$\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_i\hat{l}_i^{\beta-1}\right)^{-\theta}\hat{\chi}^{-1}=\left(\frac{E_i}{E_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{y}_i+\hat{\chi}^{-1}\sum_{k\in\mathscr{F}_D(i)}\left(\frac{\Xi_{ik}}{E_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{ik}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ik,1},\hat{t}_{ik,m}\right)\left(\hat{y}_i\hat{l}_i^{\beta-1}\right)^{-\theta}\left(\hat{l}_k^{\alpha+1}\hat{y}_k^{-\frac{\theta+1}{\theta}}\right)^{-\theta}+\hat{\chi}^{-1}\sum_{l\in\mathscr{F}_F(i)}\left(\frac{\Xi_{il}}{E_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{il}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ik,1},\hat{t}_{ik,m}\right)\left(\hat{y}_i\hat{l}_i^{\beta-1}\right)^{-\theta}\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

$$\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_i\hat{l}_i^{\beta-1}\right)^{-\theta}\hat{\chi}^{-1}=\left(\frac{Y_i}{Y_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{y}_i+\hat{\chi}^{-1}\sum_{k\in\mathscr{F}_D(i)}\left(\frac{\Xi_{ki}}{Y_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{ki}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ki,1},\hat{t}_{ki,m}\right)\left(\hat{y}_k\hat{l}_k^{\beta-1}\right)^{-\theta}\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}+\sum_{l\in\mathscr{F}_F(i)}\left(\frac{\Xi_{li}}{Y_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{li}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ki,1},\hat{t}_{ki,m}\right)\left(\hat{y}_l\right)^{-\theta}\hat{\chi}_l^{-1}\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

$$\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_l\right)^{-\theta}\hat{\chi}_l^{-1}=\left(\frac{E_l}{E_l+\sum_{k\in\mathscr{F}_D(l)}\Xi_{lk}}\right)\hat{y}_l+\sum_{k\in\mathscr{F}_D(l)}\left(\frac{\Xi_{lk}}{E_l+\sum_{k\in\mathscr{F}_D(l)}\Xi_{lk}}\right)\hat{t}_{lk}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{lk,1},\hat{t}_{lk,m}\right)\left(\hat{y}_l\right)^{-\theta}\hat{\chi}_l^{-1}\left(\hat{l}_k^{\alpha+1}\hat{y}_k^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

$$\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_l\right)^{-\theta}\hat{\chi}_l^{-1}=\left(\frac{Y_l}{Y_l+\sum_{k\in\mathscr{F}_D(l)}\Xi_{kl}}\right)\hat{y}_l+\hat{\chi}^{-1}\sum_{k\in\mathscr{F}_D(l)}\left(\frac{\Xi_{kl}}{Y_l+\sum_{k\in\mathscr{F}_D(l)}\Xi_{kl}}\right)\hat{t}_{kl}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{kl,1},\hat{t}_{kl,m}\right)\left(\hat{y}_k\hat{l}_k^{\beta-1}\right)^{-\theta}\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

Simplifying,

$$\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_i\hat{l}_i^{\beta-1}\right)^{-\theta} = \hat{\chi}\left(\frac{E_i}{E_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{y}_i + \sum_{k\in\mathscr{F}_D(i)}\left(\frac{\Xi_{ik}}{E_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{ik}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ik,1},\hat{t}_{ik,m}\right)\left(\hat{y}_i\hat{l}_i^{\beta-1}\right)^{-\theta}\left(\hat{l}_k^{\alpha+1}\hat{y}_k^{-\frac{\theta+1}{\theta}}\right)^{-\theta} + \sum_{l\in\mathscr{F}_F(i)}\left(\frac{\Xi_{il}}{E_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{il}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ik,1},\hat{t}_{ik,m}\right)\left(\hat{y}_i\hat{l}_i^{\beta-1}\right)^{-\theta}\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

$$\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}\left(\hat{y}_i\hat{l}_i^{\beta-1}\right)^{-\theta} = \hat{\chi}\left(\frac{Y_i}{Y_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{y}_i + \sum_{k\in\mathscr{F}_D(i)}\left(\frac{\Xi_{ki}}{Y_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{ki}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ki,1},\hat{t}_{ki,m}\right)\left(\hat{y}_k\hat{l}_k^{\beta-1}\right)^{-\theta}\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta} + \sum_{l\in\mathscr{F}_F(i)}\left(\frac{\Xi_{li}}{Y_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{li}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ki,1},\hat{t}_{ki,m}\right)\left(\frac{\hat{\chi}}{\hat{\chi}_l}\right)(\hat{y}_l)^{-\theta}\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

$$\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}(\hat{y}_l)^{-\theta} = \hat{\chi}_l\left(\frac{E_l}{E_l+\sum_{k\in\mathscr{F}_D(l)}\Xi_{lk}}\right)\hat{y}_l + \sum_{k\in\mathscr{F}_D(l)}\left(\frac{\Xi_{lk}}{E_l+\sum_{k\in\mathscr{F}_D(l)}\Xi_{lk}}\right)\hat{t}_{lk}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{lk,1},\hat{t}_{lk,m}\right)(\hat{y}_l)^{-\theta}\left(\hat{l}_k^{\alpha+1}\hat{y}_k^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

$$\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}(\hat{y}_l)^{-\theta}\hat{\chi}_l^{-1} = \left(\frac{Y_l}{Y_l+\sum_{k\in\mathscr{F}_D(l)}\Xi_{kl}}\right)\hat{y}_l + \hat{\chi}^{-1}\sum_{k\in\mathscr{F}_D(l)}\left(\frac{\Xi_{kl}}{Y_l+\sum_{k\in\mathscr{F}_D(l)}\Xi_{kl}}\right)\hat{t}_{kl}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{kl,1},\hat{t}_{kl,m}\right)\left(\hat{y}_k\hat{l}_k^{\beta-1}\right)^{-\theta}\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

Simplifying on the LHS and isolating the i specific terms on the second summand on the RHS,

$$\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{-\theta} = \hat{\chi}\left(\frac{E_i}{E_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{y}_i\left(\hat{y}_i\hat{l}_i^{\beta-1}\right)^{\theta} + \sum_{k\in\mathscr{F}_D(i)}\left(\frac{\Xi_{ik}}{E_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{ik}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ik,1},\hat{t}_{ik,m}\right)\left(\hat{l}_k^{\alpha+1}\hat{y}_k^{-\frac{\theta+1}{\theta}}\right)^{-\theta} + \sum_{l\in\mathscr{F}_F(i)}\left(\frac{\Xi_{il}}{E_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ik}}\right)\hat{t}_{il}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ik,1},\hat{t}_{ik,m}\right)\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

$$\left(\hat{y}_i\hat{l}_i^{\beta-1}\right)^{-\theta} = \hat{\chi}\left(\frac{Y_i}{Y_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{y}_i\left(\hat{l}_i^{\alpha+1}\hat{y}_i^{-\frac{\theta+1}{\theta}}\right)^{\theta} + \sum_{k\in\mathscr{F}_D(i)}\left(\frac{\Xi_{ki}}{Y_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{ki}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ki,1},\hat{t}_{ki,m}\right)\left(\hat{y}_k\hat{l}_k^{\beta-1}\right)^{-\theta} + \sum_{l\in\mathscr{F}_F(i)}\left(\frac{\Xi_{li}}{Y_i+\sum_{k\in\mathscr{F}(i)}\Xi_{ki}}\right)\hat{t}_{li}^{-\theta}\left(\hat{\Pi},\hat{P},\hat{t}_{ki,1},\hat{t}_{ki,m}\right)\left(\frac{\hat{\chi}}{\hat{\chi}_l}\right)(\hat{y}_l)^{-\theta}$$

Similarly for foreign locations,

$$\left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{-\theta} = \hat{\chi}_l \left(\frac{E_l}{E_l + \sum_{k\in\mathscr{F}_D(l)} \Xi_{lk}}\right) \hat{y}_l \left(\hat{y}_l\right)^{\theta} + \sum_{k\in\mathscr{F}_D(l)} \left(\frac{\Xi_{lk}}{E_l + \sum_{k\in\mathscr{F}_D(l)} \Xi_{lk}}\right) \hat{t}_{lk}^{-\theta} \left(\hat{\Pi}, \hat{P}, \hat{t}_{lk,1}, \hat{t}_{lk,m}\right) \left(\hat{l}_k^{\alpha+1} \hat{y}_k^{-\frac{\theta+1}{\theta}}\right)^{-\theta}$$

$$\left(\hat{y}_l \hat{W}_l^{-1}\right)^{-\theta} \hat{\chi}_l^{-1} = \left(\frac{Y_l}{Y_l + \sum_{k\in\mathscr{F}_D(l)} \Xi_{kl}}\right) \hat{y}_l \left(\hat{y}_l^{-\frac{\theta+1}{\theta}}\right)^{\theta} + \hat{\chi}^{-1} \sum_{k\in\mathscr{F}_D(l)} \left(\frac{\Xi_{kl}}{Y_l + \sum_{k\in\mathscr{F}_D(l)} \Xi_{kl}}\right) \hat{t}_{kl}^{-\theta} \left(\hat{\Pi}, \hat{P}, \hat{t}_{kl,1}, \hat{t}_{kl,m}\right) \left(\hat{y}_k \hat{l}_k^{\beta-1}\right)^{-\theta}$$

This system of equation pins down the equilibrium with foreign locations:

$$\hat{y}_i^{1+\theta} \hat{l}_i^{-\theta(\alpha+1)} = \hat{\chi} \left(\frac{E_i}{E_i + \sum_{k\in\mathscr{F}(i)} \Xi_{ik}}\right) \hat{y}_i^{1+\theta} \hat{l}_i^{\theta(\beta-1)} + \sum_{k\in\mathscr{F}_D(i)} \left(\frac{\Xi_{ik}}{E_i + \sum_{k\in\mathscr{F}(i)} \Xi_{ik}}\right) \hat{t}_{ik}^{-\theta} \left(\hat{\Pi}, \hat{P}, \hat{t}_{ik,1}, \hat{t}_{ik,m}\right) \hat{y}_k^{1+\theta} \hat{l}_k^{-\theta(1+\alpha)} + \sum_{l\in\mathscr{F}_F(i)} \left(\frac{\Xi_{il}}{E_i + \sum_{k\in\mathscr{F}(i)} \Xi_{ik}}\right) \hat{t}_{il}^{-\theta} \left(\hat{\Pi}, \hat{P}, \hat{t}_{ik,1}, \hat{t}_{ik,m}\right) \hat{y}_l^{1+\theta}$$

$$\hat{y}_i^{-\theta} \hat{l}_i^{\theta(1-\beta)} = \hat{\chi} \left(\frac{Y_i}{Y_i + \sum_{k\in\mathscr{F}(i)} \Xi_{ki}}\right) \hat{y}_i^{-\theta} \hat{l}_i^{\theta(\alpha+1)} + \sum_{k\in\mathscr{F}_D(i)} \left(\frac{\Xi_{ki}}{Y_i + \sum_{k\in\mathscr{F}(i)} \Xi_{ki}}\right) \hat{t}_{ki}^{-\theta} \left(\hat{\Pi}, \hat{P}, \hat{t}_{ki,1}, \hat{t}_{ki,m}\right) \hat{y}_k^{-\theta} \hat{l}_k^{\theta(1-\beta)} + \hat{\chi} \sum_{l\in\mathscr{F}_F(i)} \left(\frac{\Xi_{li}}{Y_i + \sum_{k\in\mathscr{F}(i)} \Xi_{ki}}\right) \hat{t}_{li}^{-\theta} \left(\hat{\Pi}, \hat{P}, \hat{t}_{ki,1}, \hat{t}_{ki,m}\right) \hat{\chi}_l^{-1} \hat{y}_l^{-\theta}$$

$$\hat{y}_l^{1+\theta} = \left(\frac{E_l}{E_l + \sum_{k\in\mathscr{F}_D(l)} \Xi_{lk}}\right) \hat{y}_l^{1+\theta} \hat{\chi}_l + \sum_{k\in\mathscr{F}_D(l)} \left(\frac{\Xi_{lk}}{E_l + \sum_{k\in\mathscr{F}_D(l)} \Xi_{lk}}\right) \hat{t}_{lk}^{-\theta} \left(\hat{\Pi}, \hat{P}, \hat{t}_{lk,1}, \hat{t}_{lk,m}\right) \hat{y}_k^{1+\theta} \hat{l}_k^{-\theta(1+\alpha)}$$

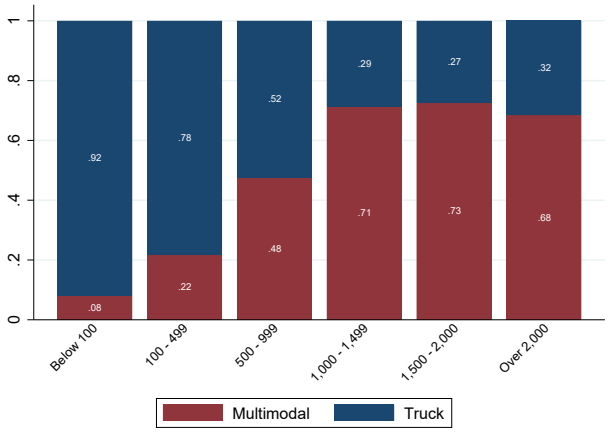
$$\hat{y}_l^{-\theta} \hat{\chi}_l^{-1} = \left(\frac{Y_l}{Y_l + \sum_{k\in\mathscr{F}_D(l)} \Xi_{kl}}\right) \hat{y}_l^{-\theta} + \hat{\chi}^{-1} \sum_{k\in\mathscr{F}_D(l)} \left(\frac{\Xi_{kl}}{Y_l + \sum_{k\in\mathscr{F}_D(l)} \Xi_{kl}}\right) \hat{t}_{kl}^{-\theta} \left(\hat{\Pi}, \hat{P}, \hat{t}_{kl,1}, \hat{t}_{kl,m}\right) \hat{y}_k^{-\theta} \hat{l}_k^{\theta(1-\beta)}$$

B Data and Additional Results

B.1 Data Appendix

Figure A.1 captures the weight share of goods that are transported within the United States by transport mode and distance.

Figure A.1. US Transport Mode Weight Shares by Distance



Notes: This figure plots the observed weight share of cargo transported by different modes across various distances in 2018. Multimodal indicates cargo movement that involves more than one mode, including truck, rail, and waterways. Source: Freight Analysis Framework, US Department of Transportation, and authors' calculations.

B.1.1 Rail Data

Our confidential carload waybills data comes from the Surface Transportation Board. This is a stratified sample of carload waybills for all U.S. rail traffic submitted by those rail carriers terminating 4,500 or more revenue carloads annually, covering 48 states (except Alaska and Hawaii). The waybills data report the origin rail station, destination rail station, and the interchange stations in between that these freight cargo are transported through. The rich geographical information in this confidential data set allows us to study the routing of these commodities through the railroad network. Additionally, this data set also contains commodity-specific information including number of car loads, weight, whether it is a domestic or international shipment, and its inter-modality—primarily indicating if the freight movement involved other transport modes, which is almost entirely containers. The intermodal container rail traffic is the fastest growing segment of rail traffic, having grown by more than 5 times between 1984 and 2019 (Figure A.2).

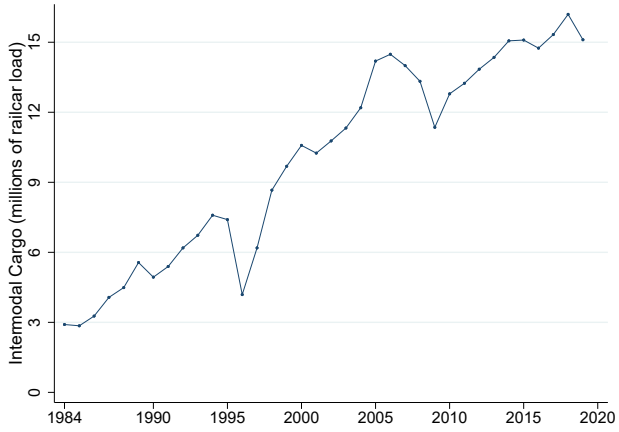
B.1.2 Road Data

We follow Allen and Arkolakis (2022) for the construction of road traffic flows data. We summarize their procedure in three steps: First, they create a sparse graph representation of the underlying road network by collapsing the high-dimensional geo-spatial information contained in the original shapefiles and only preserving nodes that are either endpoints or intersections. Furthermore, core-based statistical areas (CBSAs) are represented by a singular node along the network. Their resulting graph consists of 228 nodes and 704 edges. Second, they construct a weighted graph by including traffic data. To do so they obtain the average annual daily traffic from the 2012 Highway Performance Monitoring System (HPMS) dataset by the Federal Highway Administration and allocate it to individual links by constructing a length-weighted average of the annual daily traffic.

B.1.3 Ports

While we include the top container ports in the US in our analysis, we merge some of these ports into the one polygon due to the USACE-provided polygons either overlapping or being very close to each other. Specifically, Tacoma + Seattle are merged into a single polygon due to their port alliance (NWSA). Additionally, Tacoma and Seattle share a port alliance. Tampa + Manatee were merged do to overlap issues between the rather large Tampa region and the manually applied Manatee port

Figure A.2. Intermodal Container Rail Traffic, 1984-2019



Notes: This figure plots the observed levels of intermodal rail cargo from 1984 to 2019. Source: Confidential Carload Waybill, Surface Transportation Board, and authors' calculations.

statistical area. Los Angeles and Long Beach are also merged due to proximity. Lastly, we include Chester in the Philadelphia PSA because the USACE-provided polygon area includes Chester. Between the 28 coastal and inland ports, we capture 98 percent of total US container volumes.

B.1.4 Automatic Identification System (AIS) Vessel Traffic Data

We utilize automatic identification system (AIS) vessel traffic data from Marine Cadastre, a joint initiative between the Bureau of Ocean Energy Management and the National Oceanic and Atmospheric Administration. Here we highlight two examples to show how we capture these ships and the time they spend at a port. Panel (A) Figure A.3 shows the path of containership CMA CGM Christophe Colomb as it enters the Port of Los Angeles (LA) on May 2, 2022. It is a containership with a cargo capacity of 86,100 tons (13,800 twenty-foot equivalent unit containers (TEUs)) and is operated by container shipping company CMA CGM. Panel (B) Figure A.3 shows the path of containership Guthorm Maersk entering and leaving the Port of Newark. Guthorm Maersk is a containership with a cargo capacity of 57,000 tons (11,000 TEUs) and is operated by container shipping company Maersk. The ship path entering the port is highlighted in the figure and the redder color indicates slower speed. The darker region of both figures indicate the port polygon for both ports as defined by the US Army Corps of Engineers.

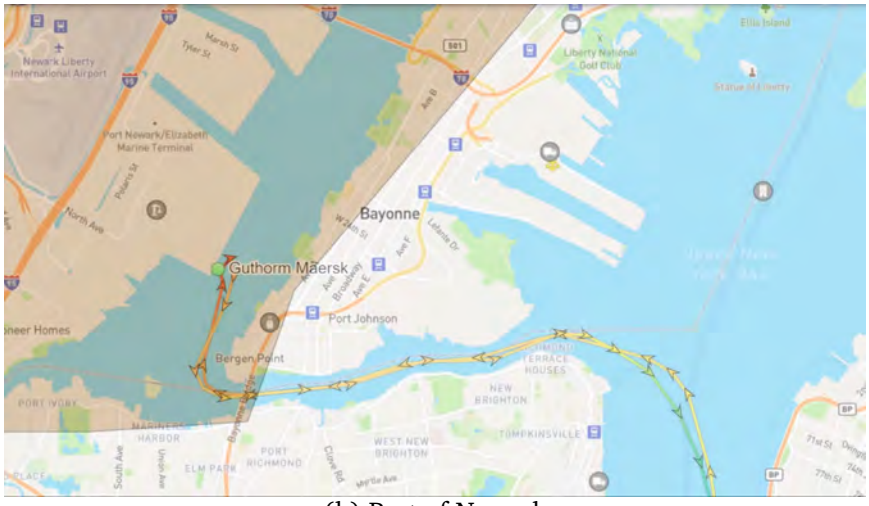
B.1.5 Matching Rail Traffic Data to Metropolitan Statistical Area (MSAs)

Using 1999 Metropolitan Statistical Area (MSA) polygons from the Census Bureau, we match the rail stations from the Waybill data to the 228 MSAs in Duranton and Turner (2011). By rail destination, we observe 224 MSAs and by origin we observe 223 MSAs. Since some of these unobserved MSAs overlap, we have 7 MSAs. We conduct our analysis on the remaining 221 MSAs, so that both the origin and destination results are comparable. The 7 MSAs that we do not have rail traffic for Daytona Beach FL, Fort Myers-Cape Coral FL, Fort Walton Beach FL, New Haven-Bridgeport-Stamford-Waterbury-Danbury CT, Providence-Warwick-Pawtucket RI, Punta Gorda FL, and Santa Fe NM.

To match the MSA-level truck vehicle-kilometers traveled (VKT) measure in Duranton and Turner (2011), we calculate the rail equivalent in two ways. First, we utilize the number of rail carloads, transported in and out of MSAs, multiplied by the weighted average of their distance traveled. We call this railcar-kilometers-traveled (rail VKT) and distinguish by destination for rail shipments transported into these MSAs, and by origin for shipments transported out of these MSAs. We also observe the weight of these rail shipments and can calculate rail weight VKT using the same method outlined previously.

Figure A.3. Illustration of AIS Mooring Paths

(a) Port of Los Angeles



(b) Port of Newark

Notes: Panel (a) shows the containership CMA CGM Christophe Colomb at the Port of Los Angeles while Panel (b) shows the containership Guthorm Maersk at the Port of Newark. The path of each ship to and from the port shows its exact travel path. The darker regions at each port shows the port polygons as defined by the US Army Corps of Engineers.

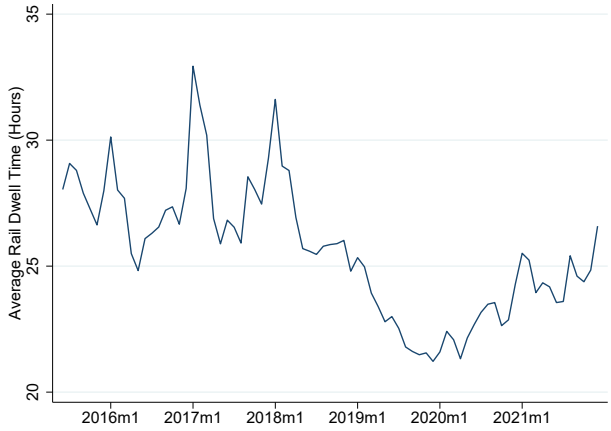
B.1.6 Rail Dwell Times Data

We obtain weekly rail station dwell times from the Surface Transportation Board (STB). Railroads provide the STB with the average time a railcar resides at a station, measured in hours, for their 10 largest stations in terms of railcars processed. This dwell time measure excludes cars on through trains—trains that travels without stops en route. Since this dataset only captures a subset of all rail stations (albeit the largest ones), we match the ports in the previous section to their local rail stations. We do this by expanding the port polygon areas in 50km intervals. The rail stations that are captured in the buffer areas of their closest port is be considered a rail station in the vicinity of this port and is likely to service traffic to and from the port. Due to their proximity, The ports of Los Angeles and Long Beach and combined into one port for this exercise. We use a buffer area of 150km which captures 7 ports and 12 rail stations. We test the robustness of this buffer area by increasing the interval in our analysis to 200km where we capture 8 ports and 14 rail stations. Further increases to this interval result in more muted responses of rail station dwell times to port traffic, as these rail stations are much further away.

Additionally, the rail dwell times dataset is reported at the weekly level. In order to match this to our daily port traffic measure for analysis, we aggregate our port traffic measure up to the weekly level. We start our week on a Monday since we observe in our data that most ships tend to enter a port on Mondays.

Summary Statistics Figure A.4 plots the average of rail station dwell times from June 2015 to December 2021. The average dwell time over this period is around 25.5 hours per station with a standard deviation of 2.5 hours. However, there is also a large decrease in dwell times around the start of the pandemic followed up a steep increase afterwards.

Figure A.4. Rail Station Dwell Times



Notes: This figure plots the average time a railcar spends at a rail station from June 2015 to July 2022.

B.2 Regression design for Modal Substitution (Subsection 4.1)

We are interested in deriving a regression that studies the impact of changes in the transport network on the modal mix in order to identify the elasticity of substitution between modes. In Subsection 4.1 we examine the results of a regression that examines location (i.e. MSA) specific outcomes as a function of local changes in the primary network. Conceptually, the regression can be seen understood to be of the following form:

$$d \ln Y_k = \alpha + \beta \times d \ln \bar{t}_{k,1} + \epsilon_{kk,1} \tag{27}$$

where Y_k refers to a local outcome (primary network traffic, or the ratio between primary and non-primary traffic) in a MSA located at node k and $d \ln \bar{t}_k$ refers to changes in transportation cost of the localized primary network.In what follows we motivate this regression by deriving a structural interpretation. In what follows, we will make an additional assumption that there exists some localized primary network fully contained within the MSA that any unimodal route originating or terminating in k needs to transition through before accessing the national primary road network. Let this localized network be represented by the transportation cost $\bar{t}_k$. In order to derive the regression we start by examining modal traffic (Equation (9)), which we restate here for convenience,

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

where we can characterize total outgoing traffic on the primary and non-primary network as,

$$\Xi_{kk,1} = P_k^{-\theta} \times t_{k,1}^{-\eta} \times \sum_{l \in \mathscr{F}_1(k)} t_{kl,1}^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta}$$

$$\Xi_{kk,m} = P_k^{-\theta} \times s_{kk,m}^{-\eta} \times \sum_{l \in \mathscr{F}_m(k)} \left(\tau_{kl,m} s_{ll,m}\right)^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta}$$

Totally differentiating with respect to the change in local transport cost on the primary network,

$$d \ln \Xi_{kk,1} = -\theta d \ln P_k - \eta d \ln t_{k,1} + d \ln \left(\sum_{l \in \mathscr{F}_1(k)} t_{kl,1}^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta} \right) \tag{28}$$

where this capture the impact of highway changes on highway traffic, plus market access changes

$$d \ln \Xi_{kk,m} = -\theta d \ln P_k + d \ln \left(\sum_{l \in \mathscr{F}_m(k)} \left(s_{kk,m} \tau_{kl,m} s_{ll,m}\right)^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta} \right)$$

which expresses changes in modal traffic as a function of the local price index ($d \ln P_k$), changes in local transport cost, either for the primary mode ($d \ln t_{k,1}$), and a final summation term that captures changes in mode-specific market access. While we could in principle regress directly on primary mode traffic as in Equation (28), the obvious challenge would be the confounding impact of changes in the local price index which will be affected by the local network changes. Instead, we propose to form the ratio between primary and non-primary modal traffic to difference out the local price index. Forming this ratio and taking the total differential, we obtain,

$$d \ln \frac{\Xi_{kk,1}}{\Xi_{kk,m}} = -\eta d \ln t_{k,1} + d \ln \left(\sum_{l \in \mathscr{F}_1(k)} \frac{t_{kl,1}^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta}}{\sum_{l \in \mathscr{F}_m(k)} \left(s_{kk,m} \tau_{kl,m} s_{ll,m}\right)^{-\eta} \times t_{kl}^{\eta-\theta} \times \Pi_l^{-\theta}} \right) \tag{29}$$

where instead now the only confounder is the final summation term that summarizes changes in the relative mode-specific market access. Lining this up with the initial motivating regression design (Equation (27)), we can see that the coefficient identifies the modal elasticity of substitution ($\beta = \eta$), while the identifying restriction is that the instrument needs to be orthogonal to changes in the relative mode-specific market access.

B.3 Elasticity of Modal Substitution: Additional Results

Since we have a slightly smaller set of MSAs due to the matching process between MSAs and our rail traffic data (see Section B.1.5 for more information), we first show that we are able to replicate the truck traffic use results from Duranton and Turner (2011) in Table A.2 (first stage results are in Table A.3). Table A.4 presents the rail traffic use results while Table A.5 presents the alternative weight-based measure of the relative truck to rail traffic use. Table A.6 presents additional robustness checks on our baseline results.

Table. A.2. Elasticity of Truck Traffic Use with respect to Road Infrastructure Improvements

	(1)	(2)	(3)	(4)	(5)
Truck Traffic Use	OLS	OLS	IV	IV	IV
Interstate Highway Lane KM	1.606	1.616	1.746	2.083	2.099
	(0.328)	(0.338)	(0.427)	(0.483)	(0.530)
Population		0.967	-0.278	-0.615	-0.484
		(0.550)	(0.303)	(0.376)	(0.393)
Geography				✓	✓
Census Divisions				✓	✓
Socioeconomic Characteristics		✓			✓
MSA FE	✓	✓			
Year FE	✓	✓	✓	✓	✓
Observations	663	663	663	663	663
R-squared	0.77	0.78	0.49	0.49	0.51
KP F-stat			13.48	10.08	10.02

Notes: Robust standard errors clustered by MSAs in parentheses. All variables are in logs. Instruments are 1835 exploration routes, 1898 railroad route kilometers, and 1947 planned interstate highways. 663 observations corresponding to 221 MSAs for each regression. See Table A.3 for first-stage regressions. The results in this table are produced using data from Duranton and Turner (2011) that is matched to rail traffic data.

Table. A.3. First Stage

	(1)	(2)	(3)
1898 Railroads	0.0879	0.0939	0.119
	(0.0460)	(0.0499)	(0.0474)
1947 Planned Interstates	0.156	0.127	0.114
	(0.0332)	(0.0322)	(0.0284)
1835 Exploration Routes	0.0249	0.0268	0.0222
	(0.0117)	(0.0124)	(0.0122)
Population	0.516	0.599	0.545
	(0.0393)	(0.0481)	(0.0597)
Geography		✓	✓
Census Divisions		✓	✓
Socioeconomic Characteristics			✓
Year FE	✓	✓	✓
Observations	663	663	663
KP F-stat	13.48	10.08	10.02

***Notes*:** See Equation (18) for more details. Robust standard errors clustered by MSAs in parentheses. All variables are in logs. Instruments are 1835 exploration routes, 1898 railroad route kilometers, and 1947 planned interstate highways. 663 observations corresponding to 221 MSAs for each regression. The results in this table are produced using data from Duranton and Turner (2011) that is matched to rail traffic data.

Table. A.4. Elasticity of Rail Traffic Use with respect to Road Infrastructure Improvements

	(1)	(2)	(3)	(4)	(5)
Rail Traffic Use	OLS	OLS	IV	IV	IV
Interstate Highway Lane KM	-0.103	-0.0993	0.434	0.254	0.401
	(0.173)	(0.175)	(0.314)	(0.337)	(0.315)
Population		0.346	0.695	0.878	0.757
		(0.299)	(0.245)	(0.286)	(0.273)
Geography				✓	✓
Census Divisions				✓	✓
Socioeconomic Characteristics		✓			✓
MSA FE	✓	✓			
Year FE	✓	✓	✓	✓	✓
Observations	663	663	663	663	663
R-squared	0.94	0.94	0.39	0.55	0.57
KP F-stat			13.48	10.08	10.02

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by MSAs in parentheses. All variables are in logs. Rail traffic use, measured in railcar-kilometers, is constructed using confidential rail waybill data. Truck traffic use (in vehicle-kilometers) and other variables are from Duranton and Turner (2011). Instruments are 1835 exploration routes, 1898 railroad route kilometers, and 1947 planned interstate highways. 663 observations corresponding to 221 MSAs for each regression. See Table A.3 for first stage regressions.

Table. A.5. Elasticity of Modal Substitution using Weight-Based Measure: Ratio of Rail to Truck Traffic Use with respect to Road Infrastructure Improvements

	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	IV	IV	IV
Interstate Highway Lane KM	-1.473	-1.472	-0.930	-1.373	-1.203
	(0.171)	(0.172)	(0.392)	(0.403)	(0.382)
Population		-0.101	0.524	1.012	0.774
		(0.308)	(0.297)	(0.338)	(0.316)
Geography				✓	✓
Census Divisions				✓	✓
Socioeconomic Characteristics		✓			✓
MSA FE	✓	✓			
Year FE	✓	✓	✓	✓	✓
Observations	658	658	658	658	658
R-squared	0.89	0.89	-0.03	0.23	0.28
KP F-stat			14.48	10.76	10.04

Notes: Robust standard errors clustered by MSAs in parentheses. All variables are in logs. Rail traffic use, measured in rail weight-kilometers, is constructed using confidential rail waybill data. Truck traffic use and all other variables are from Duranton and Turner (2011). Instruments are 1835 exploration routes, 1898 railroad route kilometers, and 1947 planned interstate highways. 663 observations corresponding to 221 MSAs for each regression. See Table A.3 for first stage regressions.

Table. A.6. Elasticity of Modal Substitution: Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rail to Truck Traffic Use	IV	IV	IV	IV	IV	IV	IV
Interstate Highway Lane KM	-1.099	-0.999	-1.220	-1.019	-1.283	-1.593	-1.703
	(0.364)	(0.405)	(0.444)	(0.430)	(0.462)	(0.528)	(0.543)
Population	0.891	1.145	1.000	0.973	0.789	1.267	1.155
	(0.306)	(0.336)	(0.367)	(0.352)	(0.376)	(0.434)	(0.443)
Geography	✓	✓	✓	✓	✓	✓	✓
Census Divisions	✓	✓	✓	✓	✓	✓	✓
Socioeconomic Characteristics	✓	✓	✓	✓	✓	✓	✓
MSA FE							
Year FE	✓	✓	✓	✓	✓	✓	✓
Rail Measure	Carload	Carload	Carload	Weight	Weight	Carload	Weight
Without 1898 Rail IV						✓	✓
Data	Total	Incoming	Outgoing	Incoming	Outgoing	Total	Total
Observations	658	658	658	658	658	658	658
R-squared	0.27	0.24	0.31	0.21	0.34	0.20	0.20
KP F-stat	10.04	10.04	10.04	10.04	10.04	14.25	14.25

Notes: Robust standard errors clustered by MSAs in parentheses. All variables are in logs. Rail traffic use, measured in rail weight-kilometers, is constructed using confidential rail waybill data. Truck traffic use (in vehicle-kilometers) and other variables are from Duranton and Turner (2011). Instruments are 1835 exploration routes, 1898 railroad route kilometers, and 1947 planned interstate highways. These observations correspond to 221 MSAs for each regression. See Table A.3 for first stage regressions. Column (1) reproduces the baseline results from Column (5) in Table 1 for comparison. Column (2) restricts the results from Column (1) to just incoming rail traffic use while Column (3) restricts the results to outgoing rail traffic use. For detailed results for weight-based measure of total rail traffic use, see Table A.5. Column (4) uses this weight-based measure of rail traffic use but restricts the results to just incoming traffic while Column (5) restricts the results to outgoing traffic. Columns (6) and (7) reruns the 2SLS regression without the 1898 railroad instrument, for the carload- and weight-based rail traffic use measures respectively.

B.4 Estimation of Intermodal Terminal Congestion: Additional Results

Table A.7 presents robustness checks on our baseline results in Table 2.

Table. A.7. Congestion Elasticity of Port Traffic with respect to Ship Dwell Times: Robustness Checks

	(1) OLS	(2) OLS	(3) IV	(4) IV	(5) IV
Port Traffic	0.229		1.111	0.208	0.272
	(0.023)		(0.511)	(0.114)	(0.163)
Port Traffic × Before Mar 2020		0.091			
		(0.011)			
Port Traffic × After Mar 2020		0.120			
		(0.012)			
Day-Month-Year FE	✓	✓	✓	✓	✓
Port-Year FE	✓	✓	✓	✓	✓
Ship-Port FE	✓	✓	✓	✓	✓
March 2020 Period				Before	After
West Coast Ports	✓		✓		
Observations	22367	90516	22367	69917	20084
First Stage KP-F			10.01	143.42	176.34

Notes: Robust standard errors in parentheses are two-way clustered at the ship and port level. All variables are in logs. Port traffic is the 28-day moving average of total daily net tonnage at the port. Weighted by ship net tonnage. Column (1) produces results using the OLS regression from Equation (19) when restricted to only west coast ports while Column (2) produces results using the OLS regression from Equation (19) with indicator variables for pre- and post-March 2020 to highlight the pandemic period. Column (3) produces the 2SLS regression using the IV from Equation (20) when restricted to only west coast ports. Column (4) produces the 2SLS regression using the IV from Equation (20) when restricted to the period before March 2020 while Column (5) is restricted to the period after March 2020. The estimates in Columns (4) and (5) are statistically significant at the 90% confidence level.

Table A.8 presents the OLS regression from Equation (19) using shorter periods of moving averages for port traffic. With shorter periods of moving averages, the ship dwell times is still positively correlated with the average tonnage at the port but the magnitudes of the coefficients are smaller.

Table. A.8. Ship Dwell Times and Port Traffic by Time Aggregations

	(1)	(2)	(3)	(4)
Port Traffic	0.0995	0.0800	0.0540	0.0246
	(0.0104)	(0.00850)	(0.0151)	(0.00692)
Day-Month-Year FE	✓	✓	✓	✓
Port-Year FE	✓	✓	✓	✓
Ship-PortFE	✓	✓	✓	✓
Moving Average (Days)	28	21	14	7
Observations	90516	90516	90515	90492
R^2	0.77	0.77	0.77	0.77
F	92.19	88.54	12.82	12.68

Notes: Robust standard errors in parentheses are two-way clustered by ship and port. All variables are in logs. Column (1) estimates the elasticity using the 28-day moving average of total daily net tonnage at the port and is replicated from the baseline results in Column (2) Table 2. Column (2) presents the 21-day moving average, Column (3) presents the 14-day moving average, and Column (4) presents the 7-day moving average. Weighted by ship net tonnage.

B.5 Rail Intermodal Terminal Congestion and Port Traffic

In this subsection, we study how port congestion can impact the multimodal network. In particular, we focus on how port traffic affects the amount of time a rail car spends at the rail station that is local to that port. We estimate the following regression (Columns (3) and (4), Table A.9):

$$\ln \text{Rail Dwell Time}_{rpwy} = \beta_2 \ln \text{Port Traffic}_{pwy} + \gamma_{wy} + \phi_{rp} + \epsilon_{rpwy} \tag{30}$$

where Rail Dwell Time$_{rpwy}$ is the average number of hours a rail car spends at a rail station r that is in the vicinity of port p during week w in year y, Port Traffic$_{pwy}$ is the average amount of port traffic at port p for that same week w in year y,[3] Port traffic is measured in both net tons of ships as in our congestion elasticity approach, as well as the number of ships. γ_{wy} is week-year fixed effects, and ϕ_{rp} is rail station-port fixed effects. The key parameter of interest, β_2, captures the link between rail dwell times and its nearest port traffic. Standard errors are clustered at the port level.

The week-year fixed effects control for aggregate events that affects all rail stations. The rail-port fixed effects control for fixed characteristics at the rail-port level. These include time-invariant comparative advantage differences across ports that result in larger capacity trains servicing the rail stations close to these ports which mechanically take longer time to unload. It also includes fixed rail station characteristics and fixed port characteristics that take into account their geography. In order to match the local rail stations in our rail dwell times dataset to ports, we extend the port areas in order to capture nearby rail stations. The buffer area we used in our baseline result is 150km which captures 6 ports and 11 rail stations. This small set of stations and ports is due to data restrictions from the rail dwell times dataset. The rail companies are only required to report the dwell times for the top 10 largest of their rail stations, and these rail stations sometimes overlap in geographic location.

We find that a one percent increase in port traffic, measured in average net tonnage, is correlated with a statistically significant increase in rail dwell times by 0.05 percent (Column (3), Table A.9). This elasticity is robust to specifications with rail station fixed effects and port fixed effects separately (Column (1), Table A.9). This elasticity is also robust to an alternative measures of port traffic by using the average number of ships (Columns (4) and (2), Table A.9).

As a robustness check, we extend the buffer area around the ports to 200km which captures 9 ports and 16 rail stations. We find that our estimate has the same sign and is within one standard error of our baseline estimate (Column (5), Table A.9). However, the magnitude of this estimate is smaller and noisy, due to the impact of port traffic being more muted on rail stations that are further away. Subsequent increases to the buffer area correspondingly result in even smaller estimates. Due to the small sample size, we are unable causally identify the multimodal impact of port traffic on rail dwell times, but we are able to show that there are statistically significant and positive correlations between the two.

[3] This measure, as mentioned from the previous subsection, is at the daily level. In order to match the rail dwell time dataset, we aggregate it up to the weekly level. We start our week on a Monday since we observe in our data that most ships tend to enter a port on Mondays.

Table. A.9. Link between Rail Dwell Times and Nearest Port Traffic

	(1)	(2)	(3)	(4)	(5)
Nearest Port Traffic (Net Tons)	0.0457		0.0457		0.0282
	(0.0152)		(0.0152)		(0.0196)
Nearest Port Traffic (Ships)		0.0372		0.0372	
		(0.0143)		(0.0143)	
Port Buffer Area	150km	150km	150km	150km	200km
Week-Month-Year FE	✓	✓	✓	✓	✓
Rail Station-Port FE					✓
Rail Station FE	✓	✓			
Observations	3327	3327	3327	3327	4316
R^2	0.80	0.80	0.80	0.80	0.79
F	9.08	6.80	9.08	6.80	2.06

Notes: Robust standard errors in parentheses are two way clustered by port and rail station. All variables are in logs. Local rail stations are determined by a 150km or 200km buffer area around the ports as described in Appendix Section B.1.6.

B.6 Model Fit: Additional Results

This section presents further results that evaluate our model's ability in representing actual observed trade flows in the data across various transport modes. Table A.10 reports the univariate regressions between residualized observed trade flows and residualized model-predicted trade flows at both the aggregate and mode-specific levels (Columns (1) and (2) respectively). Columns (3) and (4) in Table A.10 demonstrates the gravity model implications our model by comparing the negative relationship between residualized observed trade and distance versus residualized model-predicted trade and distance.

Table. A.10. Model Fit

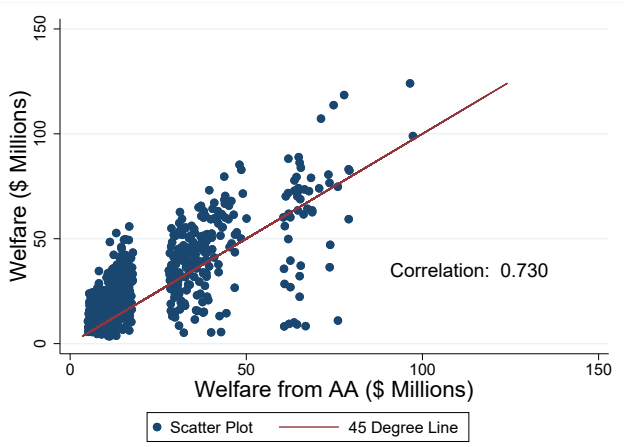
	(1)	(2)	(3)	(4)
	Residualized Observed Trade	Residualized Observed Trade	Residualized Observed Trade	Residualized Predicted Trade
Residualized Predicted Trade	1.22***			
	(0.03)			
Residualized Predicted Trade for Truck		1.20***		
		(0.03)		
Residualized Predicted Trade for Rail		1.34***		
		(0.07)		
Residualized Predicted Trade for Barges		1.02***		
		(0.26)		
Residualized Route Distance			-1.15***	-0.61***
			(0.07)	(0.02)
Conditional on origin and destination FE	✓	✓	✓	✓
Observations	14514	14514	14467	14467
R^2	0.62	0.62	0.60	0.41
F	1524.65	521.92	314.11	671.98

***Notes*:** This table compares the observed bilateral origin to destination mode-specific trade flows with the mode-specific trade flows predicted by the multimodal economic geography model based on observed traffic data along the transport network. Both the observed and predicted trade flow measures are in logs and residualized using origin and destination fixed effects, allowing for the comparison to come from similarities at the origin-destination pair-level. Column (1) compares the aggregated observed and predicted trade flows while Column (2) compares the mode-specific observed and predicted trade flows for all three transport modes (truck, rail, and barges). Columns (3) and (4) examine the gravity model implications of the model by comparing the relationship between the observed trade flows and distance (Column (3)) to the relationship between the predicted trade flows and distance (Column (4)). The route distance measure is also in logs and residualized using origin and destination fixed effects. Robust standard errors in parentheses are clustered by origin and destination cities. Weighted by trade weight in tons.

B.7 Gains Highway Link Investment Counterfactuals

In this subsection we replicate the same highway link exercise as in Allen and Arkolakis (2022). Figure A.5 shows a high correlation of 0.76 between our welfare estimates and the estimates from Allen and Arkolakis (2022). However, we estimate higher welfare effects for 81.5 percent of the links.

Figure A.5. Welfare Effects of Improving



Notes: This figure presents a comparison of the welfare impacts from adding adding ten additional lane-miles to each highway segment in the U.S., using a unimodal transport network versus a multimodal transport network. Top and bottom one percent of the data has been excluded.

B.8 Gains from Improving Intermodal Terminals Counterfactuals: Additional Results

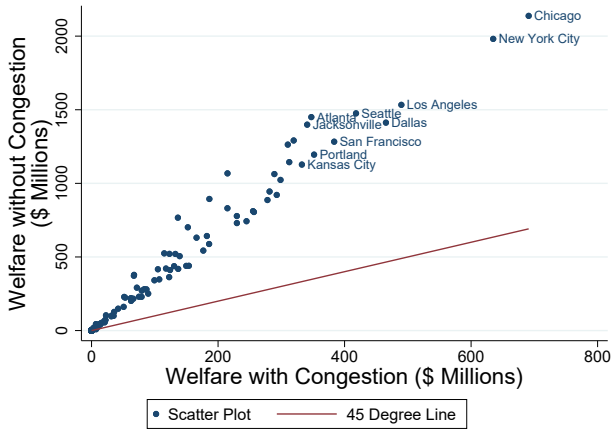
Table A.11 presents the top 30 list of intermodal terminals with the highest welfare impacts from lowering the transshipment cost in each terminal by 1 percent. Figure A.6 is a scatter plot comparing the welfare impacts from a 1 percent transshipment cost decrease at each terminal with and without congestion effects.

Table. A.11. Top 30 Ranking: Welfare Benefits of Improving Intermodal Terminals

	(1) CBSA Name	(2) Population	(3) Terminals	(4) Throughput	(5) ROI	(6) Benefit ($m)	(7) Cost ($m)
1	Chicago	9368268	88	3456228	0.249	691	553
2	New York City	14745610	29	497852	1.820	635	225
3	Los Angeles	9639715	38	2278880	-0.624	490	1301
4	Dallas	4513776	13	564160	4.581	465	83
5	Seattle	2189215	20	644052	1.917	418	143
6	San Francisco	3863536	14	104312	2.757	384	102
7	Portland	1641801	30	141432	8.891	352	36
8	Atlanta	1627623	28	610280	3.237	347	82
9	Jacksonville	936317	18	265960	6.995	341	43
10	Kansas City	1767872	55	362920	7.225	333	40
11	Memphis	997862	28	472040	6.385	320	43
12	Virginia Beach	1457869	11	173840	7.454	313	37
13	Stockton	901910	11	288320	4.876	310	53
14	St. Louis	1901086	38	197440	13.038	299	21
15	Philadelphia	4532390	33	99800	17.443	293	16
16	Salt Lake City	1198360	5	87432	11.420	289	23
17	Miami	3184615	13	96120	2.208	281	88
18	Columbus	1902095	18	176600	12.261	278	21
19	Detroit	2732964	31	185920	13.479	257	18
20	Cleveland	2175988	23	143480	15.637	255	15
21	Houston	3133212	27	210100	8.064	245	27
22	Boston	3714540	13	58720	23.948	230	9
23	Cincinnati	1744673	23	85720	22.048	230	10
24	Laredo	265111	3	43640	38.249	215	5
25	Harrisburg	655561	1	264640	5.002	215	36
26	Deltona	411646	1	35280	19.714	186	9
27	Buffalo	1150833	17	42200	44.058	186	4
28	Charlotte	1502267	20	117480	27.506	182	6
29	Minneapolis	2886766	26	134560	5.070	177	29
30	Toledo	710352	16	39040	65.768	166	2

Notes: The table shows the thirty terminals where a one percent reduction of the switching cost generates the highest benefit. Column (1) indicates the core based statistical areas (CBSA) name of the node, which includes both metropolitan and micropolitan areas. The terminal's population, number of terminals, and rail throughput in TEUs are reported in Columns (2), (3), and (4) respectively. Column (5) shows the imputed return on investment (ROI), Column (6) calculates how much 2012 US GDP would need to increase in order to match the overall welfare gain, while Column (6) presents the required cost of making this one percent cost decrease.

Figure A.6. Welfare Benefits with and without Congestion



Notes: This figure presents a comparison of the welfare impacts from a 1 percent transshipment cost decrease at each terminal with and without congestion effects. See Figure 7b for the same comparison for the top 20 most impactful terminals.